

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:38:11 ; Search time 17.7526 Seconds  
(without alignments)  
444.049 Million cell updates/sec

Title: us-09-768-826-47\_copy\_24\_105  
Perfect score: 427  
Sequence: 1 EDNDEFFMDFLQTLVGTPE.....VQVLGSQDAGTGYKDDDK 82

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR-73:  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	67.9	94	2 S17449	probable ligand-bi
2	85.5	20.0	96	2 A36581	polychlorinated bi
3	79.5	18.6	91	2 J50036	Clara cell 10K pro
4	76.5	17.9	95	2 S68231	PHG22 protein prec
5	74.5	17.4	91	1 UGRB	uteroglobin precu
6	69.5	16.3	96	1 UGMS	probable polyketid
7	69.5	16.3	2458	2 T17420	fibronectin-bindin
8	69	16.2	540	2 G86790	t123.15 protein -
9	68.5	16.0	2100	2 T38128	cell specific 10K
10	67.5	15.8	113	2 JC2026	BN11 protein - yea
11	67.5	15.8	1953	2 G3244	hypothetical prote
12	67	15.7	270	2 S30947	mannose-6-phosphat
13	67	15.7	423	2 UGRBL	uteroglobin precu
14	66.5	15.6	91	1 UGRBL	iron(III)-binding
15	66.5	15.6	358	2 A3580	probable surface prot
16	66	15.5	192	2 S70285	probable maltose/m
17	65.5	15.3	362	2 A71243	probable RNA-bind
18	65.5	15.3	609	2 T38656	phosphodiesterase
19	65.5	15.3	1687	2 T30244	exoribonuclease, v
20	65.5	15.3	1706	2 T30175	exoribonuclease, v
21	65.5	15.3	1719	2 T30174	MG331 homolog P01
22	65	15.2	209	2 S73690	repressor mode hom
23	65	15.2	255	2 B64175	hypothetical prote
24	65	15.2	312	2 G71146	hypothetical prote
25	65	15.2	842	2 E96641	coA-ligase / coenz
26	64.5	15.1	339	2 E30436	hypothetical prote
27	64.5	15.1	403	2 B98908	tyrosine transamin
28	64.5	15.1	454	1 XNRTY	H+-transporting AT
29	64	15.0	342	1 C69395	

30 64 15.0 627 2 H96951 fusion, PTS system  
31 64 15.0 6713 2 B89921 hypothetical prote  
32 63.5 14.9 1418 2 T15232 hypothetical prote  
33 63.5 14.9 1939 2 T18372 repeat organellar  
34 63 14.8 1776 2 B64549 conserved hypotet  
35 63 14.8 454 2 S69017 probable membrane  
36 63 14.8 3187 2 JC5837 364K Golgi complex  
37 62.5 14.6 258 1 OYECHF imidazoleglycerol-  
38 62.5 14.6 258 2 B90982 imidazoleglycerol-  
39 62.5 14.6 258 2 H85827 sugar transport AT  
40 62.5 14.6 362 2 A75208 unknown protein F5  
41 62.5 14.6 884 2 D96730 outer surface prot  
42 62 14.5 210 2 S69923 long-chain-fatty-a  
43 62 14.5 744 2 A54901 probable phosphono  
44 61.5 14.4 410 2 E75190 transmembrane effi  
45 61.5 14.4 1040 2 D81379

ALIGNMENTS

RESULT 1

S17449  
probable ligand-binding protein RYD5 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995-#sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S17449  
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.  
EMBO J. 10, 2813-2819, 1991  
A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac  
A:Reference number: S17447; MUID:92007724; PMID:1915264  
A:Accession: S17449  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-94 <DEA>  
A:Cross-references: EMBL:X60661; NID:957735; PIDN:CAA43068.1; PID:957736

Query Match 67.9% Score 290; DB 2; Length 94;  
Best Local Similarity 80.0% Pred. No. 1.le-24;  
Matches 56; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDNDEFFMDFLQTLVGTPEELGKYNVEDAKAAMTELKSCIDGLOPMHKAELVK 60  
Db 23 EDDNEFFMDFLQTLVGTPEELGKYNVEDAKAAMTELKSCIDGLOPMHKAELVK 82

Qy 61 LLVQVLGSQD 70  
Db 83 LLVQVLDAQE 92

RESULT 2

A36581  
polychlorinated biphenyl-binding protein precursor - rat  
A:Alternate names: Clara cell 10K secretory protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 20-Aug-1999  
C:Accession: A36581; S10185; S21676  
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta  
J. Biol. Chem. 265, 12690-12693, 1990  
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina  
A:Reference number: A36581; MUID:90324266; PMID:2115524  
A:Accession: A36581  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-96 <NOR>  
A:Cross-references: GB:J05536; NID:9206039; PIDN:AAA41817.1; PID:9206040  
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.  
Nucleic Acids Res. 18, 2939-2946, 1990

A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region  
A:Reference number: S10185; MUID:90272398; PMID:2349092  
A:Accession: S10185  
A:Status: translation not shown  
A:Molecule type: DNA

[illegible]

A:Cross-references: GB:M27564; NID:g165792; PIDN:AAA31496.1; PID:g165793  
A:Experimental source: lung  
R:Atger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.  
Biochem. J. 177, 985-988, 1979  
A:Title: N-terminal sequences of uteroglobin and its precursor.  
A:Reference number: A90303; MUID:79187160; PMID:571719  
A:Accession: A90303  
A:Molecule type: protein  
A:Residues: 1-5, 'F', 7-10, 'X', 15, 'G', 17-54, 'X', 56, 'B', 58-66, 'B', 68-70, 'XX', 73 <ATG>  
R:Ponstingl, H.; Nicot, A.; Beato, M.  
Biochemistry 17, 3908-3912, 1978  
A:Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.  
A:Reference number: A90417; MUID:79042086; PMID:568483  
A:Accession: A90417  
A:Molecule type: protein  
A:Residues: 22-81, 'Q', 83-91 <PON>  
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.  
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978  
A:Title: Amino acid sequence of a progesterone-binding protein.  
A:Reference number: A93824; MUID:79074850; PMID:281700  
A:Accession: A93824  
A:Molecule type: protein  
A:Residues: 22-49, 'D', 51, 'EN', 54-59, 61-66, 'NEPSL', 72-91 <POP>  
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.  
submitted to the Atlas, October 1982  
A:Reference number: A94608  
A:Accession: A94608  
A:Molecule type: protein  
A:Residues: 50-62:67-71 <PO2>  
R:Morize, I.; Surcouf, E.; Vaney, M.C.; Buehner, M.; Mornon, J.P.  
submitted to the Brookhaven Protein Data Bank, April 1989  
A:Reference number: A50025; PDB:1UTG  
A:Contents: annotation: X-ray crystallography, 1.34 angstroms, residues 22-91  
R:Morize, I.; Surcouf, E.; Vaney, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgrom, J.  
J. Mol. Biol. 194, 725-739, 1987  
A:Title: Refinement of the C222-1 crystal form of oxidized uteroglobin at 1.34 angstroms  
A:Reference number: A44652; MUID:88011213; PMID:3656405  
A:Contents: annotation: X-ray crystallography, 1.34 angstroms  
R:Bally, R.; Delettre, J.  
submitted to the Brookhaven Protein Data Bank, May 1989  
A:Reference number: A50553; PDB:2UTG  
A:Contents: annotation: X-ray crystallography, 1.64 angstroms, residues 22-91  
R:Bally, R.; Delettre, J.  
J. Mol. Biol. 206, 153-170, 1989  
A:Title: Structure and refinement of the oxidized p2-1 form of uteroglobin at 1.64 angstroms  
A:Reference number: A44653; MUID:89199637; PMID:2704039  
A:Contents: annotation: X-ray crystallography, 1.64 angstroms; disulfide bonds  
R:Menne, C.; Suske, G.; Arneemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982  
A:Title: Isolation and structure of the gene for the progesterone-inducible protein uteroglobin  
A:Reference number: 146904; MUID:83014990; PMID:6956897  
A:Accession: 146904  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-91 <MEN>  
A:Cross-references: GB:J00689; NID:g165786; PIDN:AAA31495.1; PID:g165788  
R:Chandra, T.; Woo, S.L.C.; Bullock, D.W.  
Biochem. Biophys. Res. Commun. 95, 197-204, 1980  
A:Title: Cloning of the rabbit uteroglobin structural gene.  
A:Reference number: 146905; MUID:81021016; PMID:7417250  
A:Accession: 146905  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 39-77 <CH2>  
A:Cross-references: GB:M25057; NID:g165802; PIDN:AAA31498.1; PID:g165803  
R:Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.  
Prog. Clin. Biol. Res. 85, 139-146, 1982  
A:Title: Structure and regulated expression of the uteroglobin gene.  
A:Reference number: 146907; MUID:82275176; PMID:6287481  
A:Accession: 146907  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-45, 'V', 47-91 <SU2>

A:Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809  
R:Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.  
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980  
A:Title: Bacterial cloning of the rabbit uteroglobin structural gene.  
A:Reference number: 146906; MUID:80241888; PMID:6156676  
A:Accession: 146906  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 53-66, 'NT', 69-72 <AT2>  
A:Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805  
C:Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It b  
C:Genetics:  
A:Introns: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: homodimer; steroid binding; uterus  
F:1-21/Domain: signal sequence #status experimental <SIG>  
F:22-91/Product: uteroglobin #status experimental <MAT>  
F:24/Disulfide bonds: Interchain (to 90) #status experimental  
F:90/Disulfide bonds: Interchain (to 24) #status experimental  
Query Match 17.4%; Score 74.5; DB 1; Length 91;  
Best Local Similarity 27.4%; Pred. No. 0.43; Mismatches 17; Indels 1; Gaps 1;  
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;  
QY 7 FMDFLOTLVGTPEELVGTGLGKYNVNEADAKAAMTELKSCDGLQPMHKAELVLLVQL 66  
DB 27 FAHVNIENLLGTPTSS-YETSLKEFEPDPTMKDAGQMCKVLDLSLPQTREINIMKTEKIV 85  
QY 67 GS 68  
DB 86 KS 87  
RESULT 6  
UGMS uteroglobin precursor - mouse  
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory prote  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence, revision 21-Jan-1997 #text\_change 22-Jun-1999  
C:Accession: A53025; A56656; I51925; S24783  
R:Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.  
Genomics 20, 27-35, 1994  
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.  
A:Reference number: A53025; MUID:94292183; PMID:8020953  
A:Accession: A53025  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <STR>  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.  
Exp. Lung Res. 19, 67-75, 1993  
A:Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecu  
A:Reference number: A56656; MUID:93178380; PMID:8440203  
A:Accession: A56656  
A:Molecule type: mRNA; protein  
A:Residues: 1-96 <SIN>  
A:Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691  
A:Experimental source: lung  
A:Note: sequence extracted from NCBI backbone (NCBI:126148)  
A:Note: parts of this sequence, including the amino end of the mature protein, were c  
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Nawkins, H.K.; DeMayo, F.J.  
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993  
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell  
A:Reference number: I51925; MUID:94000840; PMID:8398159  
A:Accession: I51925  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-96 <RES>  
A:Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093  
C:Genetics:  
A:Introns: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds

C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 16.3%; Score 69.5; DB 1; Length 96;  
Best Local Similarity 30.6%; Pred. No. 1.6;  
Matches 19; Conservative 14; Mismatches 28; Indels 1; Gaps 1;

QY 7 FMDFLQTLVGTPEELVECTLGKYNVNDKAAAMTELKSCIDGLQPMHKAELVLLVQVL 66  
DB 27 FLQVLEALLMES-ESGYVASLKPFNGSDLNAGTOLKRLVDTLPQETRIINIMKLEKIL 85  
QY 67 GS 68  
DB 86 TS 87

RESULT 7  
T17420  
probable polyketide synthase type I - Pseudomonas fluorescens  
C:Species: pseudomonas fluorescens  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000  
C:Accession: T17420  
R:Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Loper, J.E.  
J. Bacteriol. 181, 2166-2174, 1999  
A:Title: Characterization of the pyoluteorin biosynthetic gene cluster of Pseudomonas fl  
A:Reference number: 218776; MUID:99194726; PMID:10094695  
A:Accession: T17420  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2458 <NOW>  
A:Cross-references: EMBL:AF081920; NID:g4582974; PID:g2781416; PIDN:AAAC38075.1  
C:Genetics:  
A:Gene: plbB  
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier prote  
C:Keywords: carrier protein  
F:31-429/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
F:535-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>  
F:939-1009/Domain: acyl carrier protein homology <ACPI>  
F:1053-1446/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>  
F:2337-2408/Domain: acyl carrier protein homology <ACP2>

Query Match 16.3%; Score 69.5; DB 2; Length 2458;  
Best Local Similarity 32.1%; Pred. No. 68;  
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;

QY 21 ELYECTLGKYNVNDKAAAMTELKSCIDGLQPMHKAELVLLVQVLSQDGAG 73  
DB 2239 QLLHLTQSELEQDA-AARTAMOG--SGLPQRSQIVQAIARVLGGQCG 2288

RESULT 8  
G86790  
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: G86790  
R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrli  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A:Reference number: A86825; MUID:21235186; PMID:11337471  
A:Accession: G86790  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-540 <STO>  
A:Cross-references: GB:AE005176; PID:g12724308; PIDN:AAK05425.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: yngB

Query Match 16.2%; Score 69; DB 2; Length 540;  
Best Local Similarity 30.6%; Pred. No. 14;  
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 2 DNDEFFM---DFLQTLVGTPEELVECTLGKYNVNDKAAAMTELKSCIDGLQPMHKAEL 57  
DB 294 DNAEIFROKGLLNTFLNQVPNDKTSVTLENYTNEPIETALNPALSPVQNAQRYFHRQY 353  
QY 58 LVKLLVQVLSGQ 69  
DB 354 KLQNAVKFLGEO 365

RESULT 9  
T38128  
t7123.15 protein - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T38128  
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1997  
A:Reference number: 221772  
A:Accession: T38128  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2100 <BAD>  
A:Cross-references: EMBL:Z95334; PIDN:CA808600.1; GSPDB:GN00066; SPDB:SPAC2068.06  
A:Experimental source: strain 972h-; cosmid c2068  
C:Genetics:  
A:Gene: SPDB:SPAC2068.06  
A:Map position: 1

Query Match 16.0%; Score 68.5; DB 2; Length 2100;  
Best Local Similarity 33.3%; Pred. No. 73;  
Matches 21; Conservative 12; Mismatches 25; Indels 5; Gaps 2;

QY 10 FLQTLVGTPEELVECTLGKYNVNDKAAAMTELKSCIDGLQPMHKAELVLLVQVLG 67  
DB 1865 FAQLKGRVLPVTRAPLINSVSASLEKFAALDLACFSSSKP---AEVAKLLEVYS 1921  
QY 68 SQD 70  
DB 1922 SQD 1924

RESULT 10  
JC2026  
cell specific 10K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: JC2026  
R:Ray, M.K.; Magdalen, S.; O'Malley, B.W.; DeMayo, F.J.  
Biochem. Biophys. Res. Commun. 197, 163-171, 1993  
A:Title: Cloning and characterization of the mouse clara cell specific 10 kDa protein  
A:Reference number: JC2026; MUID:94071937; PMID:7916613  
A:Accession: JC2026  
A:Molecule type: DNA  
A:Residues: 1-113 <RAY>  
C:Comment: This protein is the major secretory product of the Clara cell and binds to  
C:Superfamily: uteroglobin  
F:73/Region: ochre stop codon

Query Match 15.8%; Score 67.5; DB 2; Length 113;  
Best Local Similarity 30.6%; Pred. No. 3.2;  
Matches 19; Conservative 13; Mismatches 29; Indels 1; Gaps 1;

QY 7 FMDFLQTLVGTPEELVECTLGKYNVNDKAAAMTELKSCIDGLQPMHKAELVLLVQVL 66  
DB 44 FLQVLEALLMES-ESGYVASLKPFNGSDLNAGTOLKRLVDTLPQETRIINIMKLEKIL 102  
QY 67 GS 68  
DB 103 TS 104

```

RESULT 11
S63244
BN11 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein N0646; protein YNL271c
C:Species: Saccharomyces cerevisiae
C:Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 23-Mar-2001
C:Accession: S63244; S63245; S48523; S60909; S65111
R:Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63235
A:Accession: S63244
A:Molecule type: DNA
A:Residues: 1-1553 <SEN>
A:Cross-references: EMBL:Z71547; MIPS:YNL271c
A:Experimental source: strain S288C
R:Mesenguy, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.
submitted to the Protein Sequence Database, April 1996
A:Reference number: S63245
A:Accession: S63245
A:Molecule type: DNA
A:Residues: 987-1953 <MES>
A:Cross-references: EMBL:Z71547; MIPS:YNL271c
A:Experimental source: strain S288C
R:Fares, H.F.; Pringle, J.R.
submitted to the EMBL Data Library, April 1994
A:Description: Synthetic Lethals Of CDC12.
A:Reference number: S48524
A:Accession: S48523
A:Molecule type: DNA
A:Residues: 1-937, 'A', 939-1429, 'C', 1431-1953 <FAR>
A:Cross-references: EMBL:L311766; NID:g472524; PIDN:AAA34455.1; PID:g472525
R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.N.; Hegemann, J.M.
submitted to the EMBL Data Library, October 1995
A:Description: The sequence of a 24152 bp segment from the left arm of chromosome XIV fr
A:Reference number: S60909
A:Accession: S60909
A:Molecule type: DNA
A:Residues: 1-1553 <SE2>
A:Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63225.1; PID:g1045237
R:Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A:Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from Sa
A:Reference number: S65111; MUID:96310631; PMID:8740425
A:Accession: S65111
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1553 <SEW>
A:Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63225.1; PID:g1045237
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: SGD:BN11; SHE5; SYL39
A:Cross-references: SGD:S0005215; MIPS:YNL271c
A:Map position: 14L

Query Match 15.8%; Score 67.5; DB 2; Length 1953;
Best Local Similarity 32.9%; Pred. No. 87;
Matches 24; Conservative 6; Mismatches 20; Indels 23; Gaps 4;

QY 13 TLIVGTPPEELYECTLGRYNYNEDAKAMTELKSCIDGLQPMHKAELKLVQLVLSODG- 71
||| ||| : ||| : ||| : ||| : |||
Db 804 TILNSRPEE-----SEFNKNEGTRGNSSILNS-----SEKANIQKVL-----QDGL 843

QY 72 --AGTDYKDDDDK 82
| ||||| |
Db 844 SRAKDKYKDDSK 856

RESULT 12
S30947
hypoetical protein 1 - Azotobacter vinelandii
C:Species: Azotobacter vinelandii

```

A:Residues: 1-91 <LOP>  
A:Cross-references: GB:M25609; NID:g164246; PIDN:AAA30960.1; PID:g164247  
A:Experimental source: lung  
C:Comment: Uteroglobulin, synthesized in the uterus and lung, is secreted by the uterus up  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobulin  
C:Keywords: lung; steroid binding; uterus  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: uteroglobulin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 15.6% Score 66.5; DB 1; Length 91;  
Best Local Similarity 25.8%; Pred. No. 3.3;  
Matches 16; Conservative 17; Mismatches 28; Indels 1; Gaps 1;

QY 7 FMDLQTLVGTPEELYEGLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVL 66  
DB 27 FAHVIENTLLGTPSS-YGTSLEKFPDDAMKQKRVLDLTLPQTTRENIIRKLEKIV 85  
QY 67 GS 68  
DB 86 KS 87

RESULT 15  
AD3580  
iron(III)-binding periplasmic protein precursor BMEII0565 [imported] - Brucella melitensis  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AD3580  
R:DelVecchio, V.G.; Kapral, R.J.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AD3580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <KUR>  
A:Cross-references: GB:AR008918; PIDN:AAL53807.1; PID:g17984739; \*GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEII0565  
A:Map position: II

Query Match 15.6% Score 66.5; DB 2; Length 358;  
Best Local Similarity 31.4%; Pred. No. 16;  
Matches 22; Conservative 9; Mismatches 22; Indels 17; Gaps 3;

QY 22 LYEGTLG-KYNNEDAKAAMTELKSCIDGLQPMHKAEL-----VKLLVQVL 66  
DB 142 VYAGALGWGYNIDIFAKNLEKPCWADLLDPYSYKGEIQMANPNSSGTAYTALATLVQIM 201  
QY 67 GSQDGAGTDY 76  
DB 202 G--EDKGEDY 209

Search completed: April 28, 2003, 14:44:36  
Job time : 18.7526 secs

Seq ID NO: 47 (2-105aa)  
AC No: S17449; Database: PIR-73

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:38:11 ; Search time 22.5155 Seconds  
(without alignments)  
444.049 Million cell updates/sec

Title: US-09-768-826-47\_COPY\_2\_105  
Perfect score: 537  
Sequence: 1 KGSRALLVLTFCICRMA.....VQVLGSDGAGTDKDDDK 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR-73.\*  
1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	344.5	64.2	94	2 S17449	probable ligand-bi
2	96.5	18.0	96	2 A36581	polychlorinated bi
3	92	17.1	91	2 JS0036	Clara cell 10K pro
4	86	16.0	95	2 S68231	FHG22 protein prec
5	85	15.8	91	1 UGRB	uteroglobin precu
6	81.5	15.2	96	1 UGMS	cell specific 10K
7	78.5	14.6	113	2 JC2026	signal recognition
8	76.5	14.2	504	2 S05197	SRP 34 protein - h
9	76.5	14.2	504	2 S05198	uteroglobin precu
10	76.5	14.2	504	2 S54143	mannose-6-phosphat
11	76	14.2	91	1 UGRBL	probable polyketid
12	73	13.6	609	2 T38656	fibronectin-bindin
13	70.5	13.1	423	2 S41122	t/123.15 protein -
14	69.5	12.9	2458	2 T17420	hypothetical prote
15	69	12.8	540	2 G86790	hypothetical prote
16	68.5	12.8	2100	2 T38128	BN11 protein - yea
17	68	12.7	270	2 S30947	hypothetical prote
18	67.5	12.6	299	2 T16137	outer surface prot
19	67.5	12.6	1953	2 S63244	sarcocystatin A pr
20	67	12.5	511	2 T02269	probable maltose/m
21	66.5	12.4	322	2 C83075	3c3 20c protein -
22	66.5	12.4	358	2 AD3580	phosphodiesterase
23	66	12.3	192	2 S70285	exoribonuclease, v
24	65.5	12.2	122	2 A43644	exoribonuclease, v
25	65.5	12.2	362	2 A71243	
26	65.5	12.2	1321	2 T10929	
27	65.5	12.2	1687	2 T30244	
28	65.5	12.2	1706	2 T30175	
29	65.5	12.2	1719	2 T30174	

ALIGNMENTS

RESULT 1

S17449  
Probable ligand-binding protein RYD5 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S17449  
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.  
EMBO J. 10, 2813-2819, 1991

A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac  
A:Reference number: S17447; MUID:92007724; PMID:1915264  
A:Accession: S17449  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-94 <DE>  
A:Cross-references: EMBL:X60661; NID:g57735; PIDN:CAA43068.1; PID:g57736

Query Match 64.2%; Score 344.5; DB 2; Length 94;  
Best Local Similarity 75.0%; Pred. No. 2.1e-29;  
Matches 69; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

Qy 1 KGSRALLVLTFCICRMATGDNDEFFMFLQTLVGTPEELVGLGKYNVNDAKA 60  
Db 2 KGSSA-LLVALTVLCIGLTRAEDNDEFFMFLQTLVGTPEELVGLGKYNVNDMAK 60  
Qy 61 AMTELKSCIDGLQPMHKAELVLLVQVLGSQD 92  
Db 61 ALTELKSCIDELQPVHKEQLVKLLVQVLDAQE 92

RESULT 2

A36581  
Polychlorinated biphenyl-binding protein precursor - rat  
A:Alternate names: Clara cell 10K secretory protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 20-Aug-1999  
C:Accession: A36581; S10185; S21676  
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta  
J. Biol. Chem. 265, 12690-12693, 1990  
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina  
A:Reference number: A36581; MUID:90324266; PMID:2115524  
A:Accession: A36581  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-96 <NOR>  
A:Cross-references: GB:J05536; NID:g206039; PIDN:AAA41817.1; PID:g206040  
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.  
Nucleic Acids Res. 18, 2939-2946, 1990  
A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region  
A:Reference number: S10185; MUID:90272398; PMID:2349092  
A:Accession: S10185  
A:Status: translation not shown  
A:Molecule type: DNA

A>Title: Primary structure of rabbit lung uteroglobin as deduced from the nucleotide  
FEBS Lett. 193, 247-249, 1985



A:Reference number: A24217; MUID:86056319; PMID:2415398  
A:Accession: A24217  
A:Molecule type: mRNA  
A:Residues: 22-91 <LOP>  
A:Cross-references: GB:M27564; NID:gl65792; PIDN:AAA31496.1; PID:gl65793  
A:Experimental source: lung  
R:Alger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.  
Biochem. J. 177, 985-988, 1979  
A:Title: N-terminal sequences of uteroglobin and its precursor.  
A:Reference number: A90303; MUID:79187160; PMID:571719  
A:Accession: A90303  
A:Molecule type: protein  
A:Residues: 1-5, 'F', 7-10, 'X', 15, 'G', 17-54, 'X', 56, 'B', 58-66, 'B', 68-70, 'XX', 73 <ATG>  
R:Ponstingl, H.; Nieto, A.; Beato, M.  
Biochemistry 17, 3908-3912, 1978  
A:Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.  
A:Reference number: A90417; MUID:79042086; PMID:568483  
A:Accession: A90417  
A:Molecule type: protein  
A:Residues: 22-81, 'Q', 83-91 <PON>  
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.  
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978  
A:Title: Amino acid sequence of a progesterone-binding protein.  
A:Reference number: A93824; MUID:79074850; PMID:281700  
A:Accession: A93824  
A:Molecule type: protein  
A:Residues: 22-49, 'D', 51, 'EN', 54-59, 61-66, 'NEPSL', 72-91 <POP>  
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.  
submitted to the Atlas, October 1982  
A:Reference number: A94608  
A:Accession: A94608  
A:Molecule type: protein  
A:Residues: 50-62, 67-71 <PO2>  
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Buehner, M.; Mornon, J.P.  
submitted to the Brookhaven Protein Data Bank, April 1989  
A:Reference number: A50025; PDB:1UTG  
A:Contents: annotation; X-ray crystallography, 1.34 angstroms, residues 22-91  
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgrom, J.  
J. Mol. Biol. 194, 725-739, 1987  
A:Title: Refinement of the C223-1 crystal form of oxidized uteroglobin at 1.34 angstroms  
A:Reference number: A4652; MUID:88011213; PMID:3656405  
A:Contents: annotation; X-ray crystallography, 1.34 angstroms  
R:Bally, R.; Delettre, J.  
submitted to the Brookhaven Protein Data Bank, May 1989  
A:Reference number: A50553; PDB:2UTG  
A:Contents: annotation; X-ray crystallography, 1.64 angstroms, residues 22-91  
R:Bally, R.; Delettre, J.  
J. Mol. Biol. 206, 153-170, 1989  
A:Title: Structure and refinement of the oxidized P2-1 form of uteroglobin at 1.64 angstroms  
A:Reference number: A44653; MUID:89199637; PMID:2704039  
A:Contents: annotation; X-ray crystallography, 1.64 angstroms; disulfide bonds  
R:Menne, C.; Suske, G.; Arneemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.  
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982  
A:Title: Isolation and structure of the gene for the progesterone-inducible protein uteroglobin  
A:Reference number: I46904; MUID:83014990; PMID:6956897  
A:Accession: I46904  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-91 <MEN>  
R:Chandra, T.; Woo, S.L.C.; Bullock, D.W.  
Biochem. Biophys. Res. Commun. 95, 197-204, 1980  
A:Title: Cloning of the rabbit uteroglobin structural gene.  
A:Reference number: I46905; MUID:81021016; PMID:7417250  
A:Accession: I46905  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 39-77 <CH2>  
R:Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.  
Prog. Clin. Biol. Res. 85, 139-146, 1982  
A:Title: Structure and regulated expression of the uteroglobin gene.  
A:Reference number: I46907; MUID:82275176; PMID:6287491

A:Accession: I46907  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-45, 'V', 47-91 <SUD>  
A:Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809  
R:Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.  
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980  
A:Title: Bacterial cloning of the rabbit uteroglobin structural gene.  
A:Reference number: I46906; MUID:80241888; PMID:6156676  
A:Accession: I46906  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 53-66, 'NT', 69-72 <AT2>  
A:Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805  
C:Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It b  
C:Genetics:  
A:Introns: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: homodimer; steroid binding; uterus  
F:1-21/DNA: signal sequence #status experimental <SIG>  
F:22-91/Product: uteroglobin #status experimental <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status experimental  
F:90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 15.8%; Score 85; DB 1; Length 91;  
Best Local Similarity 27.5%; Pred. No. 0.066;  
Matches 25; Conservative 17; Mismatches 33; Indels 16; Gaps 3;

Qy 6 LLLVALLFLC-----ICRMATGDEDFMDPLQLTLVGTPPELYEGTGKYNVNEDAK 59  
| | | | | || | : : : : :  
Db 7 LALVTALLCLSPASAGCP-----FAHVIENLLGLTPTS--YETSLKEFPDDTMK 56

Qy 60 AAMTELKSCIDGLQPMHKAEVLKLVQVLGS 90  
| : : | | : : : : : : : : :  
Db 57 DAGQMKNVLDLPOTTREINMKLTEKIVKS 87

RESULT 6  
UGMS  
uteroglobin precursor - mouse  
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory prote  
C:Species: Mus musculus (house mouse)  
C>Date: 03-May-1994 #sequence.revision 21-Jan-1997 #text\_change 22-Jun-1999  
C:Accession: A53025; A56656; I51925; S24783  
R:Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.  
Genomics 20, 27-35, 1994  
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.  
A:Reference number: A53025; MUID:94292183; PMID:8020953  
A:Accession: A53025  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-96 <STR>  
A:Cross-references: GB:I24372; NID:g461147; PIDN:AAA65446.1; PID:g785054  
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.  
Exp. Lung Res. 19, 67-75, 1993  
A:Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecu  
A:Reference number: A56656; MUID:93178380; PMID:8440203  
A:Accession: A56656  
A:Molecule type: mRNA; protein  
A:Residues: 1-96 <SIN>  
A:Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691  
A:Experimental source: lung  
A>Note: sequence extracted from NCBI backbone (NCBIP:126148)  
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Hawkins, H.K.; DeMayo, F.J.  
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993  
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell  
A:Reference number: I51925; MUID:94000840; PMID:8398159  
A:Accession: I51925  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-96 <RES>

A:Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093  
C:Genetics: 19/1; 81/3  
A:Introns: linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 15.2%; Score 81.5; DB 1; Length 96;  
Best Local Similarity 26.2%; Pred. No. 0.16;  
Matches 22; Conservative 19; Mismatches 42; Indels 1; Gaps 1;

QY 7 LLVATLFCICRMATGEDNDEFMDFLQTLVGTPEELYESTLGLKYNVNEAKAAMTELK 66  
Db 5 IYIVVMSICSSASSDCPGFLQVLEALLMES-ESGYVASLKPFPNGSDLQNAQTQLK 63

QY 67 SCIDGLQPMHKAELVKLLVQVLGS 90  
Db 64 RLVDLTLPQETRINIMKLEIKLTS 87

RESULT 7  
JC2026  
cell specific 10K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: JC2026  
R:Ray, M.K.; Magdalen, S.; O'Malley, B.W.; DeMayo, F.J.  
Biochem. Biophys. Res. Commun. 197, 163-171, 1993  
A:Title: Cloning and characterization of the mouse clara cell specific 10 kDa protein gene  
A:Reference number: JC2026; MUID:94071937; PMID:7916613  
A:Accession: JC2026  
A:Molecule type: DNA  
A:Residues: 1-113 <R>  
C:Comment: This protein is the major secretory product of the Clara cell and binds to pG  
C:Superfamily: uteroglobin  
F:73/Region: ochre stop codon

Query Match 14.6%; Score 78.5; DB 2; Length 113;  
Best Local Similarity 30.7%; Pred. No. 0.41;  
Matches 27; Conservative 15; Mismatches 43; Indels 3; Gaps 2;

QY 5 ALLVATLFCICRMATGEDND--EFFMDFLQTLVGTPEELYEGTLGLKYNVNEAKAAM 62  
Db 18 SLEKLLSFCFFLPATPASSDICPGFLQVLEALLMES-ESGYVASLKPFPNGSDLXNAG 76

QY 63 TELKSCIDGLQPMHKAELVKLLVQVLGS 90  
Db 77 TOLKRLVDLTLPQETRINIMKLEIKLTS 104

RESULT 8  
S05197  
signal recognition particle 54K protein - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 02-Feb-2001  
C:Accession: S05197; S10619  
R:Roemisch, K.; Webb, J.; Herz, J.; Prehn, S.; Frank, R.; Vingron, M.; Dobberstein, B.  
Nature 340, 478-482, 1989  
A:Title: Homology of 54K protein of signal-recognition particle, docking protein and two  
A:Reference number: S05197; MUID:89330596; PMID:2502717  
A:Accession: S05197  
A:Molecule type: mRNA  
A:Residues: 1-504 <R>  
A:Cross-references: EMBL:X16318; NID:g928; PIDN:CAA34385.1; PID:g930  
A:Note: part of this sequence, including the amino end of the mature protein, was confir  
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.  
Nature 340, 482-486, 1989  
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit c  
A:Reference number: S05198; MUID:89330597; PMID:2502718

A:Accession: S10619  
A:Molecule type: protein  
A:Residues: 2-35; 58-70; 120-126; 141-153; 374-388; 416-426; 470-485 <BER>  
C:Superfamily: signal recognition particle 54K protein  
C:Keywords: GTP binding; nucleotide binding; p-loop  
F:1-504/Product: signal recognition particle 54K protein #status experimental <MAT>  
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 14.2%; Score 76.5; DB 2; Length 504;  
Best Local Similarity 32.4%; Pred. No. 3.6;  
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 1 KGSRALLVATLFCICRMATGEDNDEF---FMDFLQTLVGTPEELYEGTLGLKYNVN 55  
Db 255 KGGGALSAVAATKSPFIIFGTGCEHIDDFEPKTPQFISKLLCMGDIIEGLID-----KVN 308

QY 56 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDQAGTDY 98  
Db 309 ELKLDNEALIEKLKHGQFTLRDMYEQFONIMKMGPFQOILGMIPGFGTDF 359

RESULT 9  
S05198  
signal recognition particle 54K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 23-Jul-1999  
C:Accession: S05198; S14553  
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.  
Nature 340, 482-486, 1989  
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit  
A:Reference number: S05198; MUID:89330597; PMID:2502718  
A:Accession: S05198  
A:Molecule type: mRNA  
A:Residues: 1-504 <BER1>  
A:Cross-references: EMBL:X16319  
R:Bernstein, H.  
submitted to the EMBL Data Library, November 1989  
A:Reference number: S14553  
A:Accession: S14553  
A:Molecule type: mRNA  
A:Residues: 1-42, 'DV', 45-504 <BER2>  
A:Cross-references: EMBL:X16319; NID:g54193; PIDN:CAA34386.1; PID:g54194  
C:Superfamily: signal recognition particle 54K protein  
F:2-504/Product: signal recognition particle 54K protein #status predicted <MAT>

Query Match 14.2%; Score 76.5; DB 2; Length 504;  
Best Local Similarity 32.4%; Pred. No. 3.6;  
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 1 KGSRALLVATLFCICRMATGEDNDEF---FMDFLQTLVGTPEELYEGTLGLKYNVN 55  
Db 255 KGGGALSAVAATKSPFIIFGTGCEHIDDFEPKTPQFISKLLCMGDIIEGLID-----KVN 308

QY 56 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDQAGTDY 98  
Db 309 ELKLDNEALIEKLKHGQFTLRDMYEQFONIMKMGPFQOILGMIPGFGTDF 359

RESULT 10  
S54143  
SRP 54 protein - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 23-Jul-1999  
C:Accession: S54143  
R:Patel, S.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S54143  
A:Accession: S54143  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-504 <P>  
A:Cross-references: EMBL:X86373; NID:g784998; PIDN:CAA60132.1; PID:g784999  
C:Superfamily: signal recognition particle 54K protein

```

Query Match      14.2%; Score 76.5; DB 2; Length 504;
Best Local Similarity 32.4%; Pred. No. 3.6;
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

Qy 1 KGSRALILVALTFCICRMATGEDNDEF---FMDFLQTL-LVGTPEELYEGTLGKYNVN 55
   || || || || || || || || || || || || || || || || || || || ||
Db 255 KGGGALSAVAATKSPIIFICTGTEHIDDPEFPKTPQFISKLLGMGDIEGLID-----KVN 308

Qy 56 E-----DAKAAMTELKSIDGLQPMHK--AELVKL--LVQVLGSODGAGTDY 98
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 ELKLDNDNEALIEKLKHGQFTLRDMYEQFQNLKMGKPPSOILGMTPPGFGTDF 359

```

RESULT 11

UGRBL

uterooglobin precursor - brown hare

N: Alternate names: blastokinin

C: Species: *Lepus capensis* (brown hare)

C: Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999

C: Accession: A23825

R: Lopez de Haro, M.S.; Nieto, A.

Biochem. J. 235, 895-898, 1986

A: Title: Nucleotide and derived amino acid sequences of a cDNA coding for pre-uterooglobin

A: Reference-number: A23825; MUID: 86323069; PMID: 3019311

A:Accession: A23825  
A:Molecule type: mRNA  
A:Residues: 1-91 <LOP>  
A:Cross-references: GB:N25609; NTD:g164246; PIDN:AAA30960.1; PID:g164247  
A:Experimental source: lung  
C:Comment: Uterogloblin, synthesized in the uterus and lung, is secreted by the uterus up  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uterogloblin  
C:Keywords: lung; steroid binding; uterus  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: uterogloblin #status predicted <MAT>  
F:24/disulfide bonds: interchain (to 90) #status predicted  
F:90/disulfide bonds: interchain (to 24) #status predicted

Query Match	14.2%	Score	76;	DB	1;	Length	91;
Best Local Similarity	26.4%;	Pred. No.	0.59;	Mismatches	17;	Conservative	34;
Indels	16;	Gaps	3;				
QY	6	LLLVALTLC-----TCRMATGEDNDEFTMDFLQTLLVGTPEELYEGTIGKYNVNEDAK	59				
Dd	7	LALVTALLCSPASAGIC-----PGFAHVIENTLLGTPTSS-YGTSLSKEPQDDAMK	56				
QY	60	AAMTELKSCIDGLQPMHKAELVKLLVOVLGS	90				
Dd	57	DAGMQMKVLDLTLPOTTRENIKLTEKIVKS	87				

RESULT 12  
T38656  
probable RNA-binding protein - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #accession\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T38656  
R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, August 1997  
A:Reference number: Z21804  
A:Accession: T38656  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-609 <MUR>  
A:Cross-references: EMBL:AL109739; NID:el534774; PIDN:CAB52270.1; GSPDB:GN000066; SPDB:SF  
A:Experimental source: strain 972h-; cosmid c343  
C:Genetics:  
A:Gene: SPDB:SPAC343.07  
A:Map position: 1  
A:Introns: 44/1; 96/2; 325/3

Query Match	13.6%	Score 73;	DB 2;	Length 609;
-------------	-------	-----------	-------	-------------

[illegible]

RESULT 13  
S41122  
mannose-6-phosphate isomerase (EC 5.3.1.8) - human  
N:Alternate names: phosphomannose isomerase  
C:Species: Homo sapiens (man)  
C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 08-Oct-1999  
C:Accession: S41122; S38666  
R:Roudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.  
Eur. J. Biochem. 219, 415-423, 1994  
A:Title: Purification, cDNA cloning and heterologous expression of human phosphomannose  
A:Reference number: S41122; MUID:94139717; PMID:8307007

A:Accession: S41122  
A:Molecule type: mRNA  
A:Residues: 1-423 <PRO>  
A:Cross-references: EMBL:X76057; NID:g416016; PIDN:CAA53657.1; PID:g416017  
C:Genetics:  
A:Gene: GDB:MPI  
A:Cross-references: GDB:119397; OMIM:154550  
A:Map position: 15q22-15qter  
C:Superfamily: yeast mannose-6-phosphate isomerase  
C:Keywords: Intramolecular oxidoreductase; isomerase

```

Query Match      13.1%; Score 70.5; DB 2; Length 423;
Best Local Similarity 25.7%; Pred. No. 13;
Matches 27; Conservative 19; Mismatches 38; Indels 21; Gaps 5;

Qy      7  LLVALTIF-CICRMATGEDNDETFMDFLQTLVLTGTPPELYEGTLG-----KYNVNEDA 58
      : ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      138  MAIALTPFOGLGCFRPVEE-----IVTFLKKVPE--PQFLIGDEAATHLQKTMSHDS 187

Qy      59  KAAMTELKSCIDGLQPMHK-----AELVKLLIVQVLGSDGAGCTDYKD 100
      : | : | : | | | | | | | | | | | | | | | | | | | | | | |
Db      188  QAVASSLQSCFSHLMKSEKVVYEQQLNLLVKRISQQAAGNNMED 232

```

RESULT 14  
T17420  
probable polyketide synthase type I - *Pseudomonas fluorescens*  
C:Species: *Pseudomonas fluorescens*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000  
C:Accession: T17420  
R:Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Loper, J.E.  
J. Bacteriol. 181, 2166-2174, 1999  
A:Title: Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas*  
A:Reference number: Z18776; MUID:99194726; PMID:10094695  
A:Accession: T17420  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2458 <NOW>  
A:Cross-references: EMBL:AF081920; NID:g4582974; PID:g2781416; PIDN:AAC38075.1  
C:Genetics:

A; Gene: *pltB*  
 C; Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier protein  
 C; Keywords: carrier protein  
 F; 31-429/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>  
 F; 535-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>  
 F; 939-1003/Domain: acyl carrier protein homology <ACP1>  
 F; 1053-1446/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>  
 F; 2337-2408/Domain: acyl carrier protein homology <ACP2>

Query Match 12.9%; Score 69.5; DB 2; Length 2458;

Best Local Similarity 32.1%; Pred. No. 1.2e+02;  
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;  
QY 43 ELYEGTGLGKYNVNEAKAAMTELKSCIDGLQPMHRAELVKLVQVLGSDGAG 95  
DB 2239 QLLHLTQSEGLEQDA-AARTAMQG--SGLQLPQRSQIVQAIARVLGGGQCG 2288  
RESULT 15  
G86790  
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: G86790  
R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: G86790  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-540 <STO>  
A:Cross-references: GB:AF005176; PID:g12724308; PIDN:AAK05425.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: yngB

Query Match 12.8%; Score 69; DB 2; Length 540;  
Best Local Similarity 30.6%; Pred. No. 24;  
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 24 DNDEFTM---DFLQTLVLTGPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHRAE 79  
DB 294 DNAEIFRQKGLLNTFLNQVPNDKTSVTLENYVTNEPIELALNPALSPVQNAQRVFRHYQ 353  
QY 80 LVKLLVQVLGSQ 91  
DB 354 KLQAVKFLGEQ 365

Search completed: April 28, 2003, 14:44:37  
Job time : 23.5155 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: April 28, 2003, 14:38:11 ; Search time 22.732 seconds  
(without alignments)  
444.049 Million cell updates/sec

Title: US-09-768-826-47  
Perfect score: 542  
Sequence: 1 MKGSRALLLVLTFCICRM.....VQVLGSQDAGCTDKDDDDK 105

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	349.5	64.5	94	2	probable ligand-bi
2	96.5	17.8	96	2	polychlorinated bi
3	92	17.0	91	2	Clara cell 10K pro
4	91	16.8	95	2	FHG22 protein prec
5	89	16.4	91	1	uteroglobin prec
6	81.5	15.0	96	1	uteroglobin prec
7	80	14.8	91	1	uteroglobin prec
8	78.5	14.5	113	2	cell specific 10K
9	76.5	14.1	504	2	signal recognition
10	76.5	14.1	504	2	signal recognition
11	76.5	14.1	504	2	SRP 54 protein - h
12	73	13.5	609	2	probable RNA-bind
13	70.5	13.0	423	2	mannose-6-phosphat
14	69.5	12.8	2458	2	probable polyketid
15	69	12.7	540	2	fibronectin-bindin
16	68.5	12.6	2100	2	major allergen Fel
17	68	12.5	92	2	t123.15 protein -
18	68	12.5	270	2	hypothetical prote
19	67.5	12.5	299	2	hypothetical prote
20	67.5	12.5	322	2	octaprenyl-diphosp
21	67.5	12.5	1953	2	BNII protein - yea
22	67	12.4	511	2	hypothetical prote
23	66.5	12.3	358	2	iron(III)-binding
24	66	12.2	192	2	outer surface prot
25	65.5	12.1	92	2	major allergen cha
26	65.5	12.1	122	2	sarcocystatin A pr
27	65.5	12.1	362	2	probable maltose/m
28	65.5	12.1	1321	2	3C3.20c protein -
29	65.5	12.1	1687	2	phosphodiesterase

30	65.5	12.1	1706	2	T30175
31	65.5	12.1	1719	2	T30174
32	65	12.0	209	2	S73690
33	65	12.0	255	2	B64175
34	65	12.0	312	2	G71146
35	65	12.0	842	2	E96641
36	65	12.0	2233	1	ZLNZP3
37	64.5	11.9	339	2	E90436
38	64.5	11.9	403	2	B89808
39	64.5	11.9	454	1	XNRTY
40	64.5	11.9	518	2	H85440
41	64.5	11.9	882	1	WMBEU5
42	64	11.8	190	2	T20725
43	64	11.8	342	1	C69395
44	64	11.8	446	2	G81326
45	64	11.8	627	2	H96951

ALIGNMENTS

RESULT 1

SI7449  
probable ligand-binding protein RYD5 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: SI7449  
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.  
EMBO J. 10, 2813-2819, 1991  
A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac  
A:Reference number: SI7447; MUID:92007724; PMID:1915264  
A:Accession: SI7449  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-94 <DEA>  
A:Cross-references: EMBL:X60661; NID:g57735; PIDN:CRA43068.1; PID:g57736

Query Match 64.5%; Score 349.5; DB 2; Length 94;  
Best Local Similarity 75.3%; Pred. No. 7.3e-30;  
Matches 70; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY	1	MKGSRALLLVLTFCICRMATGEDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAK 60
DB	1	MKGSSA-LLVALTVCICGLTAEEDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAK 59
QY	61	AAATELKSCIDGLQPMHKAELVQLVQVLSQD 93
DB	60	AAATELKSCIDELQPVHKEQLVQLVQVLSQD 92

RESULT 2

A36581  
polychlorinated biphenyl-binding protein precursor - rat  
N:Alternate names: Clara cell 10K secretory protein  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 08-Mar-1991 #sequence\_revision 08-Mar-1991 #text\_change 20-Aug-1999  
C:Accession: A36581; S10185; S21676  
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta  
J. Biol. Chem. 265, 12690-12693, 1990  
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina  
A:Reference number: A36581; MUID:90324266; PMID:2115524  
A:Accession: A36581  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-96 <NOR>  
A:Cross-references: GB:J05536; NID:Q206039; PIDN:AAA41817.1; PID:Q206040  
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.  
Nucleic Acids Res. 18, 2939-2946, 1990  
A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region  
A:Reference number: S10185; MUID:90272398; PMID:2349092  
A:Accession: S10185  
A:Status: translation not shown  
A:Molecule type: DNA

A:Residues: 1-18 <HAG>  
A:Cross-references: EMBL:X51318; NID:g55536; PIDN:CAA35701.1; PID:g55537  
R:Umland, T.C.; Swaminathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.  
J. Mol. Biol. 224, 441-448, 1992  
A:Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 A resolution.  
A:Reference number: S21676; MUID:92219263; PMID:1560460  
A:Contents: annotation, X-ray crystallography, 3.0 angstroms  
C:Superfamily: uteroglobin  
F:24/Disulfide bonds: Interchain (to 90) #status experimental  
F:90/Disulfide bonds: Interchain (to 24) #status experimental

Query Match 17.8%; Score 96.5; DB 2; Length 96;  
Best Local Similarity 29.8%; Pred. No. 0.004;  
Matches 25; Conservative 17; Mismatches 41; Indels 1; Gaps 1;

QY 8 LLVALTFCICRMATGSDNDEFFMDLQTLVTPPELYEGTLGKYNVNEDAKAAATELK 67  
DB 5 ITIVLMISCCSSASSDPCGFLOVLEALLGS-ESNYEAAALKPFNPASDLQNAQTOLK 63  
QY 68 SCIDGLOPMHKAELVLLVQVLGS 91  
DB 64 RLVDTLPOETRINIVKLIKILTS 87

RESULT 3  
JS0036  
Clara cell 10K protein precursor - human  
N:Alternate names: urinary protein 1  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
C:Accession: JS0036; PS0309; A56890; I38397  
R:Singh, G.; Katyal, S.L.; Brown, W.B.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia  
Biochim. Biophys. Acta 950, 329-337, 1988  
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.  
A:Reference number: JS0036; MUID:89000784; PMID:3167058  
A:Accession: JS0036  
A:Molecule type: mRNA  
A:Residues: 1-91 <SIN>  
A:Cross-references: GB:X13197; NID:g23131; PIDN:CAA31584.1; PID:g23132  
A:Accession: PS0309  
A:Molecule type: protein  
R:Bernard, A.; Roels, H.; Lauwers, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De  
Clin. Chim. Acta 207, 239-249, 1992  
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and  
A:Reference number: A56890; MUID:93009001; PMID:1395029  
A:Accession: A56890  
A:Molecule type: protein  
A:Residues: 22-45 <BER>  
A:Experimental source: urine  
A:Note: sequence extracted from NCBI backbone (NCBIP:119391)  
R:Hay, J.G.; Danel, C.; Chu, C.; Crystal, R.G.  
Am. J. Physiol. 268, 565-575, 1995  
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges  
A:Reference number: I38397  
A:Accession: I38397  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-91 <RES>  
A:Cross-references: EMBL:U01101; NID:g457932; PIDN:AAA81885.1; PID:g457933  
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b  
C:Genetics:  
A:Gene: CC10  
C:Superfamily: uteroglobin  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 17.0%; Score 92; DB 2; Length 91;  
Best Local Similarity 27.8%; Pred. No. 0.012;  
Matches 27; Conservative 18; Mismatches 36; Indels 16; Gaps 3;

QY 1 MKGSALLVLTALF-----ICRMATGEDNDEFFMDLQTLVTPPELYEGTLGKYN 54  
DB 1 MKSLCLLLVILAVHVCYEANANVC-PAVLVSYSKSLFDKV-----ERFEAVLQTF 50

A:Residues: 1-18 <HAG>  
A:Cross-references: EMBL:X51318; NID:g55536; PIDN:CAA35701.1; PID:g55537  
R:Umland, T.C.; Swaminathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.  
J. Mol. Biol. 224, 441-448, 1992  
A:Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 A resolution.  
A:Reference number: S21676; MUID:92219263; PMID:1560460  
A:Contents: annotation, X-ray crystallography, 3.0 angstroms  
C:Superfamily: uteroglobin  
F:24/Disulfide bonds: Interchain (to 90) #status experimental  
F:90/Disulfide bonds: Interchain (to 24) #status experimental

Query Match 17.8%; Score 96.5; DB 2; Length 96;  
Best Local Similarity 29.8%; Pred. No. 0.004;  
Matches 25; Conservative 17; Mismatches 41; Indels 1; Gaps 1;

QY 8 LLVALTFCICRMATGSDNDEFFMDLQTLVTPPELYEGTLGKYNVNEDAKAAATELK 67  
DB 5 ITIVLMISCCSSASSDPCGFLOVLEALLGS-ESNYEAAALKPFNPASDLQNAQTOLK 63  
QY 68 SCIDGLOPMHKAELVLLVQVLGS 91  
DB 64 RLVDTLPOETRINIVKLIKILTS 87

RESULT 3  
JS0036  
Clara cell 10K protein precursor - human  
N:Alternate names: urinary protein 1  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
C:Accession: JS0036; PS0309; A56890; I38397  
R:Singh, G.; Katyal, S.L.; Brown, W.B.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia  
Biochim. Biophys. Acta 950, 329-337, 1988  
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.  
A:Reference number: JS0036; MUID:89000784; PMID:3167058  
A:Accession: JS0036  
A:Molecule type: mRNA  
A:Residues: 1-91 <SIN>  
A:Cross-references: GB:X13197; NID:g23131; PIDN:CAA31584.1; PID:g23132  
A:Accession: PS0309  
A:Molecule type: protein  
R:Bernard, A.; Roels, H.; Lauwers, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De  
Clin. Chim. Acta 207, 239-249, 1992  
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and  
A:Reference number: A56890; MUID:93009001; PMID:1395029  
A:Accession: A56890  
A:Molecule type: protein  
A:Residues: 22-45 <BER>  
A:Experimental source: urine  
A:Note: sequence extracted from NCBI backbone (NCBIP:119391)  
R:Hay, J.G.; Danel, C.; Chu, C.; Crystal, R.G.  
Am. J. Physiol. 268, 565-575, 1995  
A:Title: Human CC10 gene expression in airway epithelium and subchromosomal locus sugges  
A:Reference number: I38397  
A:Accession: I38397  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-91 <RES>  
A:Cross-references: EMBL:U01101; NID:g457932; PIDN:AAA81885.1; PID:g457933  
C:Comment: This protein consists of two identical polypeptides linked by two disulfide b  
C:Genetics:  
A:Gene: CC10  
C:Superfamily: uteroglobin  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: Clara cell 10K protein #status experimental <MAT>

Query Match 17.0%; Score 92; DB 2; Length 91;  
Best Local Similarity 27.8%; Pred. No. 0.012;  
Matches 27; Conservative 18; Mismatches 36; Indels 16; Gaps 3;

QY 1 MKGSALLVLTALF-----ICRMATGEDNDEFFMDLQTLVTPPELYEGTLGKYN 54  
DB 1 MKSLCLLLVILAVHVCYEANANVC-PAVLVSYSKSLFDKV-----ERFEAVLQTF 50

A:Residues: 1-18 <HAG>  
A:Cross-references: EMBL:X51318; NID:g55536; PIDN:CAA35701.1; PID:g55537  
R:Umland, T.C.; Swaminathan, S.; Furey, W.; Singh, G.; Pletcher, J.; Sax, M.  
J. Mol. Biol. 224, 441-448, 1992  
A:Title: Refined structure of rat Clara cell 17 kDa protein at 3.0 A resolution.  
A:Reference number: S21676; MUID:92219263; PMID:1560460  
A:Contents: annotation, X-ray crystallography, 3.0 angstroms  
C:Superfamily: uteroglobin  
F:24/Disulfide bonds: Interchain (to 90) #status experimental  
F:90/Disulfide bonds: Interchain (to 24) #status experimental

Query Match 17.8%; Score 96.5; DB 2; Length 96;  
Best Local Similarity 29.8%; Pred. No. 0.004;  
Matches 25; Conservative 17; Mismatches 41; Indels 1; Gaps 1;

QY 8 LLVALTFCICRMATGSDNDEFFMDLQTLVTPPELYEGTLGKYNVNEDAKAAATELK 67  
DB 5 ITIVLMISCCSSASSDPCGFLOVLEALLGS-ESNYEAAALKPFNPASDLQNAQTOLK 63  
QY 68 SCIDGLOPMHKAELVLLVQVLGS 91  
DB 64 RLVDTLPOETRINIVKLIKILTS 87

RESULT 3  
JS0036  
Clara cell 10K protein precursor - human  
N:Alternate names: urinary protein 1  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
C:Accession: JS0036; PS0309; A56890; I38397  
R:Singh, G.; Katyal, S.L.; Brown, W.B.; Phillips, S.; Kennedy, A.L.; Anthony, J.; Squeglia  
Biochim. Biophys. Acta 950, 329-337, 1988  
A:Title: Amino-acid and cDNA nucleotide sequences of human Clara cell 10KDa protein.  
A:Reference number: JS0036; MUID:89000784; PMID:3167058  
A:Accession: JS0036  
A:Molecule type: mRNA  
A:Residues: 1-91 <SIN>  
A:Cross-references: GB:X13197; NID:g23131; PIDN:CAA31584.1; PID:g23132  
A:Accession: PS0309  
A:Molecule type: protein  
R:Bernard, A.; Roels, H.; Lauwers, R.; Witters, R.; Gielens, C.; Soumilion, A.; Van De  
Clin. Chim. Acta 207, 239-249, 1992  
A:Title: Human urinary protein 1: evidence for identity with the Clara cell protein and  
A:Reference number: A56890; MUID:93009001; PMID:1395029  
A:Accession: A56890  
A:Molecule type: protein



A:Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093  
A:Genetics: 19/1; 81/3  
C:Introns: 19/1; 81/3  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 15.0%; Score 81.5; DB 1; Length 96;  
Best Local Similarity 26.2%; Pred. No. 0.17;  
Matches 22; Conservative 19; Mismatches 42; Indels 1; Gaps 1;

QY 8 LLVALTLCICRMATGEDNDEFMFLOTLVLTPEELVGTGKYNVNEAKAAMTELK 67  
DB 5 IITVVLISICSSASSDDICPFLQVLEALLMES-ESGYVASLKPFPNPGSDLNAGTQLK 63

QY 68 SCIDGLQPMHKAELVKLLVQVLGS 91  
DB 64 RLVDTLPQETRINIMKLEKILTS 87

RESULT 7  
UGRBL  
uteroglobin precursor - brown hare  
N:Alternate names: blastokinin  
C:Species: Lepus capensis (brown hare)  
C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999  
C:Accession: A23825  
R:Lopez de Haro, M.S.; Nieto, A.  
Biochem. J. 235, 895-898, 1986  
A:Title: Nucleotide and derived amino acid sequences of a cDNA coding for pre-uteroglobin  
A:Reference number: A23825; MUID:86323069; PMID:3019311  
A:Accession: A23825  
A:Molecule type: mRNA  
A:Residues: 1-91 <LOP>  
A:Cross-references: GB:M55609; NID:g164246; PIDN:AAA30960.1; PID:g164247  
A:Experimental source: lung  
C:Comment: Uteroglobin, synthesized in the uterus and lung, is secreted by the uterus up  
C:Complex: homodimer linked by two disulfide bonds  
C:Superfamily: uteroglobin  
C:Keywords: lung; steroid binding; uterus  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-91/Product: uteroglobin #status predicted <MAT>  
F:24/Disulfide bonds: interchain (to 90) #status predicted  
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 14.8%; Score 80; DB 1; Length 91;  
Best Local Similarity 26.8%; Pred. No. 0.23;  
Matches 26; Conservative 18; Mismatches 37; Indels 16; Gaps 3;

QY 1 MKSRALLVLTLCICRMATGEDNDEFMFLOTLVLTPEELVGTGKYNV 54  
DB 1 MKLTITLALTLCICSPASAGIC-----PGFAHVHLLGTPSS-YGTSLKBEFQ 50

QY 55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGS 91  
DB 51 PDDAMKAGMQMKVLDLTLPQTTRENIIMKLEKIVKS 87

RESULT 8  
JC2026  
cell specific 10K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 17-Mar-1999  
C:Accession: JC2026  
R:Ray, M.K.; Magdalen, S.; O'Malley, B.W.; DeMayo, F.J.  
Biochem. Biophys. Res. Commun. 197, 163-171, 1993  
A:Title: Cloning and characterization of the mouse clara cell specific 10 kDa protein ge  
A:Reference number: JC2026; MUID:94071937; PMID:7916613  
A:Accession: JC2026

A:Molecule type: DNA  
A:Residues: 1-113 <RAY>  
C:Comment: This protein is the major secretory product of the Clara cell and binds to  
C:Superfamily: uteroglobin  
F:73/Region: ochre stop codon

Query Match 14.5%; Score 78.5; DB 2; Length 113;  
Best Local Similarity 30.7%; Pred. No. 0.43;  
Matches 27; Conservative 15; Mismatches 43; Indels 3; Gaps 2;

QY 6 ALLVALTLCICRMATGEDND--EFFMDFLOTLVLTPEELVGTGKYNVNEAKAAM 63  
DB 18 SLEKLLSFCFLPATPASSDPCFLQVLEALLMES-ESGYVASLKPFPNPGSDLNAG 76

QY 64 TELKSCIDGLQPMHKAELVKLLVQVLGS 91  
DB 77 TOLKRLVDTLPQETRINIMKLEKILTS 104

RESULT 9  
S05197  
signal recognition particle 54K protein - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 02-Feb-2001  
C:Accession: S05197; S10619  
R:Roemisch, K.; Webb, J.; Herz, J.; Prehn, S.; Frank, R.; Vingron, M.; Dobberstein, B.  
Nature 340, 478-482, 1989  
A:Title: Homology of 54K protein of signal-recognition particle, docking protein and  
A:Reference number: S05197; MUID:89330596; PMID:2502717  
A:Accession: S05197  
A:Molecule type: mRNA  
A:Residues: 1-504 <ROE>  
A:Cross-references: EMBL:X16318; NID:g928; PIDN:CAA34385.1; PID:g930  
A:Note: part of this sequence, including the amino end of the mature protein, was con  
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.  
Nature 340, 482-486, 1989  
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit  
A:Reference number: S05198; MUID:89330597; PMID:2502718  
A:Accession: S10619  
A:Molecule type: protein  
A:Residues: 2-35;58-70;120-126;141-153;374-388;416-426;470-485 <BER>  
C:Superfamily: signal recognition particle 54K protein  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:1-504/Product: signal recognition particle 54K protein #status experimental <MAT>  
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 14.1%; Score 76.5; DB 2; Length 504;  
Best Local Similarity 32.4%; Pred. No. 3.7;  
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 2 KGSRALLVALTLCICRMATGEDNDEF----FMDFLQTLVLTPEELVGTGKYNV 56  
DB 255 KGGGALSAAVATKSPFIIFGTGEHIDDEPPFKTQPFISKLLGMGIDGLID-----KVN 308

QY 57 E----DAKAAMTELKSCIDGLQPMHKAELVKL--LVQVLGSQDCAGTGY 99  
DB 309 ELKLDNEALIEKLRHGQFTLRDMYEQFNIMKMGPFQILGMIPGFGTDF 359

RESULT 10  
S05198  
signal recognition particle 54K protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 23-Jul-1999  
C:Accession: S05198; S14553  
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.  
Nature 340, 482-486, 1989  
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit  
A:Reference number: S05198; MUID:89330597; PMID:2502718  
A:Accession: S05198  
A:Molecule type: mRNA  
A:Residues: 1-504 <BERL>  
A:Cross-references: EMBL:X16319





Best Local Similarity 32.1%; Pred. No. 1.2e+02;  
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;  
Qy 44 ELYEGTLGKYNVNEAKAAATELKSCIDGLQPMHKAELVQLVQLGSGDGA 96  
Db 2239 QLLHLTQSEQELEQDA-AARTAMQG--SGLQLPQRSIVQAIARVLGGCGCG 2288  
RESULT 15  
G86790  
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: G86790  
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: G86790  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-540 <STO>  
A:Cross-references: GB:AE005176; PID:gl2724308; PIDN:AAK05425.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: yngB

Query Match 12.7%; Score 69; DB 2; Length 540;  
Best Local Similarity 30.6%; Pred. No. 25;  
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;  
Qy 25 DNDEFFM---DFLOTLLVGTPEELYEGTLGKYNVNEAKAAATELKSCIDGLQPMHKA 80  
Db 294 DNAEIFRQKGLLNTFLNQVPNDKTSVTLENYTNEPIEIALNPALSPVQNAQRVHRYQ 353  
Qy 81 LVKLLVQVLGSQ 92  
Db 354 KLQAVKFLGEQ 365

Search completed: April 28, 2003, 14:44:35  
Job time : 25.732 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:39:06 ; Search time 16.9588 Seconds  
(without alignments)  
182.172 Million cell updates/sec

Title: US-09-768-826-47  
Perfect score: 542  
Sequence: 1 MKGRALLVALTLCICRM.....VOVLGSDGACTDYKDDDK 105

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCT05\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	25.2	70	US-09-620-093A-6	Sequence 6, Appli
2	92	17.0	91	US-08-455-896-8	Sequence 8, Appli
3	92	17.0	91	US-08-933-149-8	Sequence 8, Appli
4	92	17.0	91	US-09-082-343-8	Sequence 8, Appli
5	92	17.0	91	US-09-082-253-8	Sequence 8, Appli
6	92	17.0	91	PCT-US96-08235-8	Sequence 8, Appli
7	91	16.8	90	US-08-821-451A-4	Sequence 4, Appli
8	91	16.8	90	US-09-263-810-4	Sequence 4, Appli
9	91	16.8	90	US-08-912-276-15	Sequence 15, Appli
10	91	16.8	90	US-09-583-169-4	Sequence 4, Appli
11	91	16.8	90	US-09-215-818-6	Sequence 6, Appli
12	85.5	15.8	75	US-08-864-357F-3	Sequence 3, Appli
13	79.5	14.7	70	US-08-864-357F-1	Sequence 1, Appli
14	74.5	13.7	69	US-08-864-357F-2	Sequence 2, Appli
15	71	13.1	90	US-08-821-451A-2	Sequence 2, Appli
16	71	13.1	90	US-09-263-810-2	Sequence 2, Appli
17	71	13.1	90	US-09-583-169-2	Sequence 2, Appli
18	69.5	12.8	75	US-08-864-357F-4	Sequence 4, Appli
19	68	12.5	92	US-07-807-529A-2	Sequence 2, Appli
20	68	12.5	92	PCT-US93-02462-2	Sequence 2, Appli
21	68	12.5	94	US-07-662-193-1	Sequence 1, Appli
22	68	12.5	94	US-08-300-928C-2	Sequence 2, Appli
23	68	12.5	94	US-08-430-944D-2	Sequence 2, Appli
24	68	12.5	94	US-08-430-014-2	Sequence 2, Appli
25	68	12.5	94	US-08-431-184-2	Sequence 2, Appli
26	68	12.5	2233	US-08-569-853-1	Sequence 1, Appli
27	68	12.5	2233	US-08-569-853-2	Sequence 2, Appli

28	68	12.5	2233	3	US-08-987-439-1	Sequence 1, Appli
29	66.5	12.3	606	4	US-08-891-298-3	Sequence 3, Appli
30	66.5	12.3	606	4	US-08-653-648A-11	Sequence 11, Appli
31	66	12.2	69	4	US-08-912-276-23	Sequence 23, Appli
32	66	12.2	1139	4	US-09-046-158A-22	Sequence 22, Appli
33	65	12.0	139	4	US-08-559-896B-4	Sequence 4, Appli
34	65	12.0	652	4	US-08-559-896B-2	Sequence 2, Appli
35	63.5	11.7	745	1	US-08-453-472-5	Sequence 5, Appli
36	63.5	11.7	745	1	US-08-038-948-9	Sequence 9, Appli
37	63.5	11.7	745	1	US-08-453-952-5	Sequence 5, Appli
38	63.5	11.7	745	2	US-08-484-993B-43	Sequence 43, Appli
39	63.5	11.7	745	2	US-08-862-903-5	Sequence 5, Appli
40	63.5	11.7	745	2	US-08-484-158B-43	Sequence 43, Appli
41	63.5	11.7	745	2	US-08-484-596A-43	Sequence 43, Appli
42	63.5	11.7	745	2	US-08-480-150A-43	Sequence 43, Appli
43	63.5	11.7	745	3	US-08-458-731-43	Sequence 43, Appli
44	63.5	11.7	745	3	US-08-149-223A-43	Sequence 43, Appli
45	63.5	11.7	1621	1	US-08-242-677-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-620-093A-6  
; Sequence 6, Application US/09620093A  
; Patent No. 6420337  
; GENERAL INFORMATION:  
; APPLICANT: Iyer, Subramanian  
; APPLICANT: Nguyen, Tay N.  
; APPLICANT: Wu, Dauh-Rung  
; APPLICANT: Xing, Ruve  
; TITLE OF INVENTION: HIGHLY PURIFIED CYTOKINE ACTIVATING FACTOR AND METHODS  
; FILE REFERENCE: 3812-7  
; CURRENT APPLICATION NUMBER: US/09/620,093A  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/145,317  
; PRIOR FILING DATE: 1999-07-23  
; PRIOR APPLICATION NUMBER: 60/197,619  
; PRIOR FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Gallus gallus  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (54)  
; OTHER INFORMATION: xaa = any residue  
; NAME/KEY: UNSURE  
; LOCATION: (63)  
; OTHER INFORMATION: xaa = any residue  
US-09-620-093A-6

Query Match 26.2%; Score 142; DB 4; Length 70;  
Best Local Similarity 49.2%; Pred. No. 1.2e-10;  
Matches 30; Conservative 7; Mismatches 24; Indels 0; Gaps 0;  
QY 28 EFFMDFLOTLVGTPEELYEGTIGKYNVNDKAAATELKSCIDGLOPMHKAELVKLLVQ 87  
Db 6 EVVPSLIQTLLGSGTEQLVAGPISRYNVNDMTSAAALAEKKCIDELPPXHLKALVNLXKQ 65  
QY 88 V 88  
Db 66 I 66

RESULT 2  
US-08-455-896-8  
; Sequence 8, Application US/08455896  
; Patent No. 5668267





CURRENT APPLICATION DATA: US/O  
 APPLICATION NUMBER: 10/20080100  
 FILING DATE: 20080100  
 CLASSIFICATION: 717  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Becker, Cheryl L

```
;
;
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 5972. US.P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6183952e
;
; US-08-912-276-15

Query Match 16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps 2;

Qy 1 MKGSRALLVALTLC-----ICRMATGEDNDEFFMDFLQTLVLVGTPEELYEGTLGKYN 54
Db 1 MKLSVCLLLVTLALCCYQANAEFCPALVSELLDFFFS-----EPLFKLSLAKFD 50

Qy 55 VNEDAKAAMTELKSCIDGLOPMHKAELVKLLVQVL 89
Db 51 APPEAVALKGVKRCCTDMSLQKRSIAEVLVKIL 85

RESULT 10
US-09-583-169-4
; Sequence 4, Application US/09583169
; Patent No. 6338948
; GENERAL INFORMATION:
; APPLICANT: Jiao Ni, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/583,169
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,451
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J. G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
;
; US-09-583-169-4
```

```
Query Match 16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps 2;

Qy 1 MKGSRALLVALTLC-----ICRMATGEDNDEFFMDFLQTLVLVGTPEELYEGTLGKYN 54
Db 1 MKLSVCLLLVTLALCCYQANAEFCPALVSELLDFFFS-----EPLFKLSLAKFD 50

Qy 55 VNEDAKAAMTELKSCIDGLOPMHKAELVKLLVQVL 89
Db 51 APPEAVALKGVKRCCTDMSLQKRSIAEVLVKIL 85

RESULT 11
US-09-215-818-6
; Sequence 6, Application US/09215818A
; Patent No. 6379671
; GENERAL INFORMATION:
; APPLICANT: Colpitts, Tracey
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
; TITLE OF INVENTION: DETECTING DISEASES OF THE BREAST
; FILE REFERENCE: 5972.US.P2
; CURRENT APPLICATION NUMBER: US/09/215,818A
; EARLIER FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/912,276
; EARLIER FILING DATE: 1997-08-17
; EARLIER APPLICATION NUMBER: 08/697,105
; EARLIER FILING DATE: 1996-08-19
; EARLIER APPLICATION NUMBER: 08/912,149
; EARLIER FILING DATE: 1997-08-15
; EARLIER APPLICATION NUMBER: 08/697,106
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo Sapiens
;
; US-09-215-818-6

Query Match 16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps 2;

Qy 1 MKGSRALLVALTLC-----ICRMATGEDNDEFFMDFLQTLVLVGTPEELYEGTLGKYN 54
Db 1 MKLSVCLLLVTLALCCYQANAEFCPALVSELLDFFFS-----EPLFKLSLAKFD 50

Qy 55 VNEDAKAAMTELKSCIDGLOPMHKAELVKLLVQVL 89
Db 51 APPEAVALKGVKRCCTDMSLQKRSIAEVLVKIL 85

RESULT 12
US-08-864-357F-3
; Sequence 3, Application US/08864357F
; Patent No. 6255281
; GENERAL INFORMATION:
; APPLICANT: Claragen, Inc. & NIH
; TITLE OF INVENTION: Use of Recombinant Human Uteroglobin in Treatment of Inflamm
; TITLE OF INVENTION: Fibrotic Conditions
; FILE REFERENCE: 116142/2
; CURRENT APPLICATION NUMBER: US/08/864,357F
; CURRENT FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 75
; TYPE: PRT
; ORGANISM: rat
;
; US-08-864-357F-3

Query Match 15.8%; Score 85.5; DB 4; Length 75;
```

```

Db          65 KS 66

RESULT 15
US-08-821-451A-2
; Sequence 2, Application US/08821451A
; Patent No. 6066724
; GENERAL INFORMATION:
; APPLICANT: Jian NI, Guo-Liang Yu and Relner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,451A
; FILING DATE: March 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,724
; FILING DATE: March 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-821-451A-2

Query Match 13.1%; Score 71; DB 3; Length 90;
Best Local Similarity 26.3%; Pred. No. 0.14;
Matches 25; Conservative 14; Mismatches 40; Indels 16; Gaps 2;

Qy 1 MKGSRALLLVALTLFC-----ICRMATGEDNDEFFMDFLQTLTLVGTPPELYEGTLGKYN 54
Db 1 MRLSVCLLWVSLALCCYQAHVLCPPAVASEITVFLFLSDAANVL-----QVAKLN 50

Qy 55 VNEDAKAAMTELKSCIDGLQPMHKAEVLKLLVQVL 89
Db 51 PPPEALAAKLEVRHCTDQISFKKRLSLKELVLEIV 85

Search completed: April 28, 2003, 14:45:28
Job time : 18.9588 secs

```



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:36:31 ; Search time 20.2887 seconds  
(without alignments)  
832.774 Million cell updates/sec

Title: US-09-768-826-47\_COPY\_24\_105

Perfect score: 427

Sequence: 1 EDNDEFFMDFLQTLVGTPE.....VQVLGSQDGAGTDYKDDDDK 82

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_21.\*
- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362	84.8	95	4	Q8TD33
2	290	67.9	94	11	Q05702
3	86.5	20.3	478	5	O62571
4	76	17.8	2104	16	Q8XM24
5	74.5	17.4	96	11	Q8VD96
6	69.5	16.3	2458	2	O51827
7	69	16.2	540	16	O9CFY4
8	68.5	16.0	2100	3	P87112
9	68	15.9	1425	4	O95247
10	68	15.9	1483	4	O95277
11	68	15.9	1527	4	O9UIG0
12	67.5	15.8	317	10	Q8VXX7
13	67.5	15.8	454	11	Q802R1
14	67.5	15.8	485	10	Q9LVG2
15	67.5	15.8	1933	3	O13450
16	66.5	15.6	358	16	Q8YCG5

17	66	15.5	151	10	Q8RYJ9
18	66	15.5	189	2	O85808
19	66	15.5	192	2	P94238
20	66	15.5	210	2	O44978
21	66	15.5	742	6	Q95JV6
22	66	15.5	3071	12	Q8QXL2
23	66	15.5	3071	12	Q8QXL1
24	65.5	15.3	362	17	O57942
25	65.5	15.3	454	6	Q9XSW4
26	65.5	15.3	457	5	Q9VAB8
27	65.5	15.3	609	3	Q9UT83
28	65.5	15.3	795	5	O62006
29	65.5	15.3	1687	11	O35651
30	65.5	15.3	1706	11	P97790
31	65.5	15.3	1719	11	P97789
32	65	15.2	262	5	Q9BPS1
33	65	15.2	312	17	O58122
34	65	15.2	785	10	O9FRN0
35	65	15.2	842	10	O9SY89
36	64.5	15.1	287	16	Q8XN94
37	64.5	15.1	339	17	Q97VJ6
38	64.5	15.1	403	16	Q99WH0
39	64.5	15.1	454	11	Q9QWS4
40	64.5	15.1	714	10	Q9FJ70
41	64	15.0	627	16	Q9L8G6
42	64	15.0	652	5	Q9U0D4
43	64	15.0	6713	16	Q99U54
44	64	15.0	6713	16	Q931R6
45	63.5	14.9	1418	5	O01837

ALIGNMENTS

RESULT 1

Q8TD33	Q8TD33	PRELIMINARY;	PRT;	95 AA.
AC	Q8TD33;			
DT	01-JUN-2002 (TReMBLrel. 21, Created)			
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)			
DE	Putative ligand binding protein RYD5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RP	Bingle C.D.;			
RT	"Human RYD5, a new secretogloblin."			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY026938; AAK08972.1;			
SQ	SEQUENCE 95 AA; 10457 MW; F629AF06C96D2392 CRC64;			

Query Match 84.8%; Score 362; DB 4; Length 95;  
Best Local Similarity 98.6%; Pred. No. 1.4e-33;  
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	EDNDEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVK	60
DB	24	EDNDEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVK	83
QY	61	LLVQVLGSQDGA	72
DB	84	LLVQVLGSQDGA	95

RESULT 2

Q05702	Q05702	PRELIMINARY;	PRT;	94 AA.
ID	Q05702			
AC	Q05702;			
DT	01-NOV-1996 (TReMBLrel. 01, Created)			
DT	01-NOV-1996 (TReMBLrel. 01, Last sequence update)			

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Potential ligand-binding protein.
GN RYD5.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=92007724; PubMed=1915264;
RA Dear T.N., Boehm T., Keverne E.B., Rabbitts T.H.;
RT "Novel genes for potential ligand-binding proteins in subregions of
RL the olfactory mucosa.";
DR EMBO J. 10:2813-2819 (1991).
DR EMBL; X60661; CAA43068.1; -.
DR InterPro: IPR000329; Uterogloblin.
DR SMART; SM00096; UTG; 1.
SQ SEQUENCE 94 AA; 10401 MW; 25A4BBB4977E247 CRC64;

Query Match 67.9%; Score 290; DB 11; Length 94;
Best Local Similarity 80.0%; Pred. No. 2.1e-25;
Matches 56; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDNDEFMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVK 60
Db 23 EDNDEFMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVK 82
Qy 61 LLVQVLGSD 70
Db 83 LLVQVLDAQE 92

RESULT 3
ID O62571 PRELIMINARY; PRT; 478 AA.
AC O62571;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Stress-responsive protein kinase (PRKSD).
OS Suberites domuncula (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Hadromerida; Suberitidae; Suberites.
OX NCBI_TaxID=55567;
RN [1]
RP SEQUENCE FROM N.A.
RC Mueller W.E.G.;
RA Mueller W.E.G.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96394688; PubMed=8798342;
RA Kruse M., Gamulin V., Cetkovic H., Pancer Z., Mueller I.M.,
RA Mueller W.E.G.;
RT "Molecular evolution of the Metazoan protein kinase C multigene
family.";
RL J. Mol. Evol. 43:374-383 (1996).
DR EMBL; Y13101; CAA73555.1; -.
DR HSSP; P24941; IHCL.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 478 AA; 54098 MW; 6B38FF5FD8D6332 CRC64;

Query Match 20.3%; Score 86.5; DB 5; Length 478;
Best Local Similarity 30.6%; Pred. No. 0.22;
Matches 26; Conservative 18; Mismatches 24; Indels 17; Gaps 4;
```

```
Qy 1 EDNDEF---FMDFLQTLVLVGTPEELYECT-LGKYNVNEDAKAAMTELKSCIDGLQPMHKA 56
Db 240 KDTOKFSNEFSDFIISRLVKNPEERMSATALLQHKFIKSAK-PVAVLK----- 286

Qy 57 ELVKLLVQVLGSDQAGTDYKDDDD 81
Db 287 ELIADAMRILEEGSGSDEDDND 311

RESULT 4
ID O8XM24 PRELIMINARY; PRT; 2104 AA.
AC O8XM24;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable alpha-N-acetylglucosaminidase.
GN CPE0866.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001 (2002).
DR EMBL; AP003188; BAB80572.1; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR00421; FA58_C.
DR InterPro: IPR003961; FN_III.
DR Pfam; PF00754; F5_F8_type_C; 1.
DR Pfam; PF00041; fn3; 1.
DR SMART; SM00060; FN3; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 2104 AA; 236010 MW; 5FE345659599BED7 CRC64;

Query Match 17.8%; Score 76; DB 16; Length 2104;
Best Local Similarity 29.0%; Pred. No. 22;
Matches 27; Conservative 10; Mismatches 34; Indels 22; Gaps 3;

Qy 1 EDNDEF---FMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKA 57
Db 721 KDSADFVDFADILKQLLANSAQEEYEVCMNAYNNGCKEKFVSGKF-----LE 770

Qy 58 LVKLLVQVLGSDQ-----DGAGTDYKDDDD 81
Db 771 LKQLQVRLSTRPEFLIGNWIEDARTMLKSDDD 803

RESULT 5
ID Q8VD96 PRELIMINARY; PRT; 96 AA.
AC Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CC10 protein precursor.
GN CC10.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG.
RA Gutierrez-Sagal R., Nieto A.;
```

LI SEQUENCE FROM N.A.  
 RP  
 RC STRAIN-572;  
 RA Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CD -1- SIMILARITY: TO YEAST NTL.  
 DR EMBL; Z95334; CAB08600.1; -.  
 CC EMBL; Z95334; CAB08600.1; -.  
 KW Hypothetical protein; Nucleic protein; Transcription regulation.  
 SQ SEQUENCE 2100 AA; 237132 MW; 760ADA97C408E1CC CRC64;

Query Match 15.9%; Score 68; DB 4; Length 1527;  
Best Local Similarity 23.7%; Pred. No. 1.2e+02;  
Matches 23; Conservative 20; Mismatches 34; Indels

```
Oy 2 DNDEFFDFLQTLVGTPEELGKYNVNEDAKAAMTELK-----SCIDGLQPMHK 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1046 DGNQELLNFRSLDLIEVATRLKGLGLGVETSEFARVLSLEKLKDFGECVIALQ----- 1101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 56 AELVKLLVQ-----VLGSDGAGTGYKDDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1102 ASVIKKFLQGFMAPKQRRKLOSDSAKTEEVDEEK 1138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q8VXX7 ID Q8VXX7 PRELIMINARY; PRT; 317 AA.
AC Q8VXX7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative APTALAZ.
GN AT5G60120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kaniya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.; 15238620.
RT "Full length cDNA of gene At5g60120 (GI:15238620).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074373; AAL67069.1;
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR SMART; SM00380; AP2; 1.
DR SEQUENCE 317 AA; 35036 MW; F95F8FF802608C17 CRC64;

Query Match 15.8%; Score 67.5; DB 10; Length 317;
Best Local Similarity 28.8%; Pred. No. 19;
Matches 15; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

Oy 31 NVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQLGSDGAGTGYKDDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 SVTSNSVVNAEASSCIDGDELCSTRVTKQFEIL---KGGGEEEEDDDE 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
Q8QZRI ID Q8QZRI PRELIMINARY; PRT; 454 AA.
AC Q8QZRI;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to tyrosine aminotransferase (Hypothetical 50.6 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RA TISSUE-LIVER;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RA TISSUE-LIVER;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC024120; AAH24120.1;
DR EMBL; BC024264; AAH24264.1;
DR EMBL; BC025934; AAH25934.1;
KW Hypothetical protein; Aminotransferase; Transferase.
SQ SEQUENCE 454 AA; 50565 MW; CDB6E16CDB5DC7D CRC64;

Query Match 15.8%; Score 67.5; DB 11; Length 454;
Best Local Similarity 33.9%; Pred. No. 30;
Matches 19; Conservative 11; Mismatches 21; Indels 5; Gaps 2;

Oy 11 LQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQV 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 LKSLQRTPOEFYQDTLSFLKSNADLCYGAL-----SAIPGLQVPRPSGMYLMVGI 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9LVG2 ID Q9LVG2 PRELIMINARY; PRT; 485 AA.
AC Q9LVG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AP2 domain transcription factor-like.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RC MEDLINE=20181125; Pubmed=10718197;
RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB019231; BAA96941.1;
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR SMART; SM00380; AP2; 1.
DR SEQUENCE 485 AA; 53164 MW; BEFB290E407FC3D5 CRC64;

Query Match 15.8%; Score 67.5; DB 10; Length 485;
Best Local Similarity 28.8%; Pred. No. 33;
Matches 15; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

Oy 31 NVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQLGSDGAGTGYKDDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 SVTSNSVVNAEASSCIDGDELCSTRVTKQFEIL---KGGGEEEEDDDE 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
O13450 ID O13450 PRELIMINARY; PRT; 1953 AA.
AC O13450;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 'PPF3P'.
GN BNI1 OR PPF3 OR YNL271C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180;
RA Yorihuzi T., Ohsumi Y.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D38411; BAA22512.1;
DR SGD; S0005215; BNI1.
```

DR InterPro; IPR003104; FH2.  
DR Pfam; PF02181; FH2; 1.  
DR SMART; SM00498; FH2; 1.  
SQ SEQUENCE 1953 AA; 219671 MW; 0C5A9D5280A5858F CRC64;  
  
Query Match 15.8%; Score 67.5; DB 3; Length 1953;  
Best Local Similarity 32.9%; Pred. No. 1.9e+02;  
Matches 24; Conservative 6; Mismatches 20; Indels 23; Gaps 4;  
  
QY 13 TLLVGTPEELYEGTLGKYNVEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSDG- 71  
|:| ||| :||| :| | || :| | ||| :| | |||  
Db 804 TILNSRPEE-----SFKNNEGTRGMSSLNS-----SEKANIQKVL-----ODGL 843  
  
QY 72 ---AGTDYKDDDK 82  
| | ||||| |  
Db 844 SRAKKDYKDDSKK 856

Search completed: April 28, 2003, 14:43:29  
Job time : 23.2887 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:33:15 ; Search time 39.1684 Seconds  
(without alignments)  
278.964 Million cell updates/sec

Title: us-09-768-826-47\_COPY\_24\_105

Perfect score: 427

Sequence: 1 EDNDFPFMFLQTLVLTPE.....VQVLGSDGAGTDYKDDDK 82

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	427	100.0	105	22	Human secreted pro-
2	427	100.0	105	23	Human albumin fusi
3	362	84.8	95	23	Human secreted pro
4	142	33.3	70	22	Cytokine activatin
5	85.5	20.0	75	20	Rat uteroglobin-li
6	85.5	20.0	75	21	Rat uteroglobin
7	81	19.0	96	18	hPSTI, SSI chimeric
8	79.5	18.6	70	15	Lymphocytoma proli
9	79.5	18.6	70	18	Target peptide fro
10	79.5	18.6	70	20	Human uteroglobin-

11	79.5	18.6	70	21	Human uteroglobin.
12	79.5	18.6	91	22	Human Clara cell p
13	76.5	17.9	78	18	hPSTI,OMTKY3 chime
14	76.5	17.9	78	18	hPSTI,SSImutant ch
15	76.5	17.9	102	22	Bovine mammary tis
16	76.5	17.9	452	22	Novel human diagno
17	74.5	17.4	69	20	Rabbit uteroglobin
18	74.5	17.4	69	21	Rabbit uteroglobin
19	73	17.1	936	22	Murine mature chor
20	73	17.1	962	22	Murine chordin (CH
21	69.5	16.3	75	20	Mouse uteroglobin-
22	69.5	16.3	75	21	Murine uteroglobin
23	69.5	16.3	119	21	Lung cancer associ
24	69.5	16.3	187	23	Mature human LP-48
25	69	16.2	47	19	Ad5 fiber knob dom
26	68	15.9	540	23	Lactococcus lactis
27	68	15.9	1527	19	Human BAZ1-beta pr
28	68	15.9	1531	19	Human BAZ1-beta pr
29	67	15.7	167	23	Protein rcd44HS.
30	67	15.7	423	22	Human polypeptide
31	67	15.7	465	22	Human polypeptide
32	66	15.5	95	21	Borna disease viru
33	66	15.5	1139	19	Human JAK2 protein
34	65.5	15.3	457	22	Drosophila melanog
35	65	15.2	139	18	Plasmodium falci
36	65	15.2	323	22	Murine mature chor
37	65	15.2	345	22	Murine chordin-lik
38	65	15.2	652	18	Plasmodium falci
39	65	15.2	842	23	Herbicidally activ
40	64.5	15.1	92	23	Human ORF1596 prot
41	64.5	15.1	1594	23	Modified human/mou
42	64	15.0	24	22	Flag epitope tag f
43	64	15.0	2086	22	Staphylococcus aur
44	64	15.0	5795	22	Staphylococcus aur
45	64	15.0	6281	22	Staphylococcus aur

#### ALIGNMENTS

RESULT 1  
AAB70079  
ID AAB70079 standard; Protein: 105 AA.

XX AAB70079;

XX 14-MAY-2001 (first entry)

XX Human secreted protein #18.

XX Human; secreted protein; immunomodulatory; antisclerotic;  
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;  
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;  
KW antialkheimers; antiparkinsonian; antimicrobial; vulnary; gene therapy;  
KW immune disorder; hyperproliferative; cardiovascular; angiogenic;  
KW neurological; infection.

OS Homo sapiens.

XX WO200112776-A2.

XX 22-FEB-2001.

XX 15-AUG-2000; 2000WO-US22350.

XX 16-AUG-1999; 99US-0148759.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;

XX WPI; 2001-244245/25.

DR N-PSDB; AAF76860.  
XX  
XX Nucleic acids encoding 18 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
PT disease and diabetic retinopathy.  
XX  
XX Claim 11; Page 373; 380pp; English.  
XX  
XX The present sequence is one of 18 novel human secreted proteins. The  
CC nucleic acids encoding the proteins and the proteins themselves may be  
CC used in the prevention, diagnosis and treatment of diseases including  
CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus  
CC and human immunodeficiency virus (HIV) infections), hyperproliferative  
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary  
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft  
CC neovascularisation and diabetic retinopathy), neurological disorders  
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
CC infectious diseases and/or for promoting wound healing, regeneration  
CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA  
CC probes in diagnostic assays to detect and quantitate the presence of  
CC similar nucleic acid sequences in samples. The polypeptides may also be  
CC used as antigens in the production of antibodies and in assays to  
CC identify modulators of protein expression and activity.  
XX  
XX Sequence 105 AA;  
SQ  
Query Match 100.0%; Score 427; DB 22; Length 105;  
Best Local Similarity 100.0%; Pred. No. 5.7e-45;  
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 60  
DB 24 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 83  
QY 61 LLVQVLGSDGAGTDYKDDDDK 82  
DB 84 LLVQVLGSDGAGTDYKDDDDK 105  
RESULT 2  
ABG65513  
ID ABG65513 standard; Protein; 105 AA.  
XX  
XX AC ABG65513;  
XX  
XX 27-AUG-2002 (first entry)  
XX  
XX Human albumin fusion protein #2188.  
XX  
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.  
XX  
XX Homo sapiens.  
OS  
OS Synthetic.  
XX  
XX WO200177137-A1.  
XX  
XX 18-OCT-2001.  
PD  
XX  
XX 12-APR-2001; 2001WO-US11988.  
PF  
XX  
XX 12-APR-2000; 2000US-229358P.  
PR  
XX 25-APR-2000; 2000US-199384P.  
PR  
XX 21-DEC-2000; 2000US-256931P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA

XX  
PI Rosen CA, Haseltine WA;  
XX  
XX WPI; 2002-010886/01.  
XX  
XX New fusion protein for treating disease e.g. diabetes comprises an  
PT albumin fused to a therapeutic protein.  
XX  
XX Claim 1; Page 2082; 2102pp; English.  
XX  
XX The present invention relates to albumin fusion proteins comprising a  
CC therapeutic protein X and human albumin (HA, also known as human serum  
CC albumin, HSA). The proteins are useful for treating a disease or  
CC disorder that may be modulated by therapeutic protein X. The albumin  
CC extends the shelf-life of protein X, and may increase its biological  
CC in vitro/in vivo activity. The protein is useful for treating and  
CC diagnosing disorders such as cancer, reproductive disorders, digestive  
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
CC (e.g. diabetes), haematopoietic disorders, neural disorders  
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
CC (e.g. osteoporosis, arthritis). ABG6326-ABG65518 represent albumin  
CC fusion proteins of the invention.  
XX  
XX Sequence 105 AA;  
SQ  
Query Match 100.0%; Score 427; DB 23; Length 105;  
Best Local Similarity 100.0%; Pred. No. 5.7e-45;  
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 60  
DB 24 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 83  
QY 61 LLVQVLGSDGAGTDYKDDDDK 82  
DB 84 LLVQVLGSDGAGTDYKDDDDK 105  
RESULT 3  
AAO21685  
ID AAO21685 standard; Protein; 95 AA.  
XX  
XX AC AAO21685;  
XX  
XX 05-SEP-2002 (first entry)  
XX  
XX Human secreted protein SEQ ID No 27.  
XX  
XX Antiartherosclerotic; cytostatic; HIV; anti allergic; antianemic;  
KW antiasthmatic; cardiant; vasotropic; neuroprotective; nootropic; SBGP;  
KW anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;  
KW immunosuppressive; human secreted protein; cell proliferative disorder;  
KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;  
KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;  
KW ischaemic heart disease; congestive heart failure; neurological disorder;  
KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;  
KW Parkinson's disease; epilepsy; stroke; knockin humanised animal;  
KW transgenic animal; gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200238602-A2.  
XX  
XX 16-MAY-2002.  
PD  
XX  
XX 08-NOV-2001; 2001WO-US47420.  
PF  
XX  
XX 08-NOV-2000; 2000US-247505P.  
PR  
XX 09-NOV-2000; 2000US-248642P.  
PR  
XX 16-NOV-2000; 2000US-249824P.  
XX  
XX 21-NOV-2000; 2000US-252824P.  
PR







PR 25-OCT-1995; 95US-0548186.  
XX (ARRI-) ARRIS PHARM CORP.  
XX Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;  
XX WPI; 1997-259023/23.  
DR N-PSDB; AAT84524.  
XX  
PT Modified subtilisin inhibitor - having altered reactive site  
PT resulting in decreased immunogenicity, used in detergent  
PT composition(s)  
XX  
PS Example 2; Fig 9A; 69pp; English.  
XX  
CC This polypeptide comprises human pancreatic secretory trypsin  
CC inhibitor (PSII), modified to include a Streptomyces subtilisin  
CC inhibitor (SSI) reactive site. A phoA signal peptide facilitates  
CC expression in E. coli. The polypeptide is encoded by a DNA  
CC construct (AAT84524) produced by PCR amplification of overlapping  
CC synthetic oligonucleotides. Novel modified subtilisin inhibitors  
CC are claimed in which a human standard mechanism inhibitor, such as  
CC PSII, has an altered reactive site so that its dissociation  
CC constant for inhibition of subtilisin is reduced by at least a  
CC factor of 100. Also claimed are: a nucleic acid encoding the  
CC subtilisin inhibitor; expression vectors; and transformed host  
CC cells. The novel subtilisin inhibitors combine the low  
CC allergenicity of human standard mechanism inhibitors, which are  
CC not specific for subtilisin, and the high affinity of non-human  
CC subtilisin inhibitors such as SSI. They form a complex with  
CC subtilisin, so as to avoid problems of allergenicity in detergent  
CC formulations, and can also be used in affinity purification and  
CC (diagnostic) quantification of subtilisin.  
XX  
SQ Sequence 96 AA;  
  
Query Match 19.0%; Score 81; DB 18; Length 96;  
Best Local Similarity 28.1%; Pred. No. 0.033;  
Matches 25; Conservative 8; Mismatches 30; Indels 26; Gaps 3;  
  
Qy 15 LVGTPPEELVGTGKYNV-----NEDAKAAM-----TELKSCIDGLQPM 53  
Db 13 LLFTPSQAAGNTGADSLGKAKYNEVDVMPYVPCVGTGNTYNECVLCFEG 58  
Qy 54 HKAEVLKLVGVLSQDGAGDYKDDDK 82  
Db 69 -RKROTSILIOKSGPCGGGSDYKDDDK 96  
  
RESULT 8  
AAR55778  
ID AAR55778 standard; peptide; 70 AA.  
AC AAR55778;  
XX  
XX 21-FEB-1995 (first entry)  
XX  
XX Lymphocytoma proliferation activating peptide (LPAP).  
XX  
KW Lymphocytoma proliferation activating peptide; LPAP; tumour;  
KW degenerative disease; lymphocyte; mesenchyme tissue;  
KW epidermal injury; cardiovascular disease; respiratory disease;  
KW urogenital disease; nervous system; immune system; antibody;  
KW diagnosis.  
XX  
OS Homo sapiens.  
XX  
XX key Location/Qualifiers  
XX Disulfide-bond 3  
XX  
FT /note= "When the peptide exists in dimeric form the  
FT cysteine at position 3 forms a disulfide bond  
FT with the cysteine at position 69 on the  
FT antiparallel peptide strand."

FT Disulfide-bond 69  
FT /note= "When the peptide exists in dimeric form the  
FT cysteine at position 69 forms a disulfide bond  
FT with the cysteine at position 3 on the  
FT antiparallel peptide strand."  
XX  
XX DE4244565-A.  
XX  
XX 07-JUL-1994.  
XX  
XX 30-DEC-1992; 92DE-4244565.  
XX  
XX 30-DEC-1992; 92DE-4244565.  
XX (FORS/) FORSMANN W.  
XX Adermann K, Aoki A, Forssmann W, Hock D, Meyer M;  
XX Raida M, Schulz-knappe;  
XX WPI; 1994-219093/27.  
XX  
XX Lymphocytoma proliferation activating peptide structure and  
XX peptide - its prodn and its use for the treatment of, eg tumours,  
XX hypoplasia, skin-, cardiovascular-, respiratory- and  
XX urogenital-diseases  
XX  
XX Claim 1; Page 16; 38pp; German.  
XX  
XX Lymphocytoma proliferation activating peptide (LPAP) is used as a  
XX medicament for the treatment of disturbances in the cell growth of  
XX normal and tumour cells; degenerative diseases of humans e.g. of  
XX lymphocytes and other mesenchyme tissues; epidermal injuries;  
XX cardiovascular-, respiratory- and urogenital diseases; the nerve  
XX system and the immune system. Antibodies directed against LPAP can  
XX also be used in diagnostic assays.  
XX  
SQ Sequence 70 AA;  
  
Query Match 18.6%; Score 79.5; DB 15; Length 70;  
Best Local Similarity 27.4%; Pred. No. 0.034;  
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;  
  
Qy 7 FMDLQTLVGTPEELVGTGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVOYL 66  
Db 6 FQRTVETLLMDTPSS-YEAMELFSPDQDMREAGAGLKLVDTLTPQKPRESIKLMKIA 64  
Qy 67 GS 68  
Db 65 QS 66  
  
RESULT 9  
AAW11628  
ID AAW11628 standard; peptide; 70 AA.  
XX  
XX AC AAW11628;  
XX  
XX DT 07-OCT-1997 (first entry)  
XX  
XX Target peptide from an ubiquitin fusion protein.  
XX  
DE Recombinant peptide; host cell; expression vector; hormone;  
DE antimicrobial; epitope; ACTH; VIP; GHRH; CGRP; amylin; osteocalcin;  
KW insulin; natriuretic peptide.  
XX  
XX Synthetic.  
XX  
XX WO9701627-A1.  
XX  
XX PD 16-JAN-1997.  
XX  
XX PF 27-JUN-1996; 96WO-US10858.  
XX

```

PR 03-NOV-1995; 95US-0553006.
PR 27-JUN-1995; 95US-0000822.
XX (IGEN-) IGEN INC.
PA (YOST/) YOST P B.
XX
PI Lohnas GL, Pilon AL, Roberts SF, Yost PB;
XX WPI: 1997-100203/09.
XX
XX High level expression of recombinant peptide(s) - using host cells
XX comprising an expression vector encoding a ubiquitin-peptide fusion
XX protein
XX
XX Claim 40; Page 56; 72pp; English.
XX
XX A fermentation method has been discovered for producing a target peptide
XX from an ubiquitin fusion protein. The method involves inducing the
XX synthesis of the fusion protein by a transformant contained in a
XX fermentation culture; culturing the transformant under conditions
XX suitable for intracellular production and accumulation of the induced
XX fusion protein; recovering the fusion protein in at least 50% specific
XX yield; hydrolysing the induced fusion protein using an ubiquitin
XX specific hydrolase to release the target peptide; and recovering the
XX target peptide. The present sequence represents a specifically claimed
XX example of a target peptide. The method can be used to produce e.g.
XX peptide hormones, antimicrobial peptides, peptide epitope fragments,
XX ACTH, VIP, GRH, CGRP, amylin, osteocalcin, insulin, and natriuretic
XX peptides. The system can provide recombinant fusion proteins in high
XX specific yields of over 50% of total cellular protein.
XX
XX Sequence 70 AA;
SQ
Query Match 18.6%; Score 79.5; DB 18; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
Qy 7 FMDFLQTLVGTPELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQL 66
Db 6 FORVIETLLMDTPSS-YEAAWELFSPDQDMREAGAKLKLVDLTLPQKPRESIKLMEXIA 64
Qy 67 GS 68
Db 65 QS 66
RESULT 10
AAW87569
ID AAW87569 standard; protein; 70 AA.
XX
AC AAW87569;
XX
XX 12-MAR-1999 (first entry)
XX
XX Human uteroglobin-like protein CC10.
XX
XX Uteroglobin; phospholipase A2; fibronectin; inflammation; asthma;
XX cystic fibrosis; premature labour; infertility; rheumatoid arthritis;
XX type I diabetes; nephropathy; inflammatory bowel disease;
XX Crohn's disease; ulcerative colitis; pancreatitis; peritonitis; allergy;
XX multiple organ failure; adult respiratory distress syndrome;
XX acute renal failure; organ transplant rejection; autoimmune uveitis;
XX corneal transplant surgery; neonatal RDS; cytomegalovirus retinitis;
XX pneumonia; cystitis; schistosomiasis; vaginal candidiasis; fibrosis;
XX neonatal broncho-pulmonary dysplasia; haemodialysis; glomerulopathy;
XX artificial insemination.
XX
XX Homo sapiens.
XX
XX WO9853846-A1.
XX
XX 03-DEC-1998.
XX

```

---

```

PF 28-MAY-1998; 98WO-US11026.
XX
XX 28-MAY-1997; 97US-0864357.
XX
XX (CLAR-) CLARAGEN INC.
XX (USSH ) US NAT INST OF HEALTH.
XX
XX Mukherjee AB, Pilon AL, Zhang Z;
XX
XX WPI: 1999-059777/05.
XX
XX Treating and preventing inflammation and fibrosis with human
XX uteroglobin - which inhibits phospholipase A2 and binds to
XX fibronectin, for clinical or cosmetic use, e.g. in respiratory
XX distress syndrome
XX
XX Disclosure; Fig 1; 59pp; English.
XX
XX The present sequence represents a uteroglobin-like protein. Recombinant
XX human uteroglobin inhibits phospholipase A2 (PLA2), and binds to
XX fibronectin. Inhibition of PLA2 is used to treat or prevent a wide range
XX of systemic and ocular inflammations, asthma, cystic fibrosis, premature
XX labour, infertility, rheumatoid arthritis, type I diabetes, nephropathy,
XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,
XX pancreatitis, peritonitis, allergy, multiple organ failure, adult
XX respiratory distress syndrome (RDS), acute renal failure, inflammation
XX secondary to infection or surgery, and organ transplant rejection. Some
XX specified applications are in autoimmune uveitis, corneal transplant
XX surgery, neonatal and adult RDS, cytomegalovirus retinitis, pneumonia,
XX cystitis, schistosomiasis and vaginal candidiasis. Fibrotic conditions
XX that can be treated are pulmonary, renal and vascular fibrosis.
XX Uteroglobin may be administered to correct deficiency in endogenous
XX uteroglobin, e.g. in neonatal broncho-pulmonary dysplasia, complications
XX of haemodialysis and inherited glomerulopathy. Uteroglobin can also be
XX used to increase the rate of artificial insemination, in humans or
XX animals, by treatment of sperm, fertilised eggs or embryos before
XX transfer to the uterus.
XX
XX Sequence 70 AA;
SQ
Query Match 18.6%; Score 79.5; DB 20; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
Qy 7 FMDFLQTLVGTPELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQL 66
Db 6 FORVIETLLMDTPSS-YEAAWELFSPDQDMREAGAKLKLVDLTLPQKPRESIKLMEXIA 64
Qy 67 GS 68
Db 65 QS 66
RESULT 11
AAB32128
ID AAB32128 standard; protein; 70 AA.
XX
XX AAB32128;
XX
XX 14-FEB-2001 (first entry)
XX
XX Human uteroglobin.
XX
XX Human; uteroglobin; immunoglobulin A mediated disease; IgA nephropathy;
XX autoimmune disorder; pulmonary inflammation; Wegener's granulomatosis;
XX Goodpasture's disease; diabetic glomerulosclerosis.
XX
XX Homo sapiens.
XX
XX WO200062795-A2.
XX
XX 26-OCT-2000.
XX

```



```

DT 02-DEC-1997 (first entry)
DE hpSTI.OMTKY3 chimeric subtilisin inhibitor.
XX
XX Subtilisin inhibitor; protease inhibitor; human; PSTI;
KW hPSTI.OMTKY3; pancreatic secretory trypsin inhibitor;
KW turkey ovomucoid third domain protein; enzyme engineering;
KW protein engineering; detergent.
XX
XX Chimeric Homo sapiens;
OS Chimeric Weleagris gallopavo;
OS Chimeric synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..3
FT Protein /label= sig_peptide
FT Peptide /label= Mat_protein
FT Peptide 22..31
FT /label= OMTKY3
FT /note= "turkey ovomucoid third domain protein
FT Peptide 66..78
FT /label= M2 tag
FT /note= "affinity tag facilitates purification"
XX
XX WO9715670-Al.
XX
XX 01-MAY-1997.
XX
XX 25-OCT-1996; 96WO-US17153.
XX
XX 25-OCT-1995; 95US-0548186.
XX
XX (ARRI-) ARRIS PHARM CORP.
XX
XX Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
XX
XX WPI; 1997-259023/23.
XX
XX N-PSDB; AAT84525.
XX
XX Modified subtilisin inhibitor - having altered reactive site
XX resulting in decreased immunogenicity, used in detergent
XX composition(s)
XX
XX Example 2; Fig 10A; 69pp; English.
XX
XX This polypeptide comprises human pancreatic secretory trypsin
XX inhibitor (PSTI), modified to include a turkey ovomucoid third
XX domain (OMTKY3) reactive site. It is encoded by a DNA construct
XX (AAT84525) produced by PCR amplification of overlapping synthetic
XX oligonucleotides. Novel modified subtilisin inhibitors are
XX claimed in which a human standard mechanism inhibitor, such as
XX PSTI, has an altered reactive site so that its dissociation
XX constant for inhibition of subtilisin is reduced by at least a
XX factor of 100. Also claimed are: a nucleic acid encoding the
XX subtilisin inhibitor; expression vectors; and transformed host
XX cells. The novel subtilisin inhibitors combine the low
XX allergenicity of human standard mechanism inhibitors, which are
XX not specific for subtilisin, and the high affinity of non-human
XX subtilisin inhibitors such as OMTKY3. They form a complex with
XX subtilisin, so as to avoid problems of allergenicity in detergent
XX formulations, and can also be used in affinity purification and
XX (diagnostic) quantification of subtilisin.
XX
XX Sequence 78 AA;
XX
XX Query Match 17.9%; Score 76.5; DB 18; Length 78;
XX Best Local Similarity 35.4%; Pred. No. 0.092;
XX Matches 17; Conservative 5; Mismatches 21; Indels 5; Gaps 1;
XX
XX 35 DAKAAMTELKSCIDGLQPMKAEVLKLVQVLGSDGAGTDYKDDDDK 82
XX | | | | | : | | | | | | | | | | | | | | | |
XX

```

Db 36 DGNTPNECVLCFEG-----RRKQTSILIQSGPGGGGSDYKDDDDK 78

RESULT 14  
AAW26364  
ID AAW26364 standard; Protein; 78 AA.  
XX  
XX AAW26364;  
XX  
XX 02-DEC-1997 (first entry)  
XX  
XX hpSTI.SSImutant chimeric subtilisin inhibitor.  
XX  
KW Subtilisin inhibitor; protease inhibitor; human; PSTI; hpSTI.SSI;  
KW pancreatic secretory trypsin inhibitor; SSI; enzyme engineering;  
KW protein engineering; detergent.  
XX  
OS Chimeric Homo sapiens;  
OS Chimeric Streptomyces sp.;  
OS Chimeric synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..3  
FT Protein /label= sig\_peptide  
FT Peptide /label= Mat\_protein  
FT Peptide 22..32  
FT /label= SSI  
FT /note= "Streptomyces subtilisin inhibitor mutated  
FT reactive site contains Ala at P6, Ala at  
FT P5, Ala at P4, Leu at P1 and Arg at P3."  
FT Peptide 67..78  
FT /label= M2 tag  
FT /note= "affinity tag facilitates purification"  
XX  
XX WO9715670-Al.  
XX  
XX 01-MAY-1997.  
XX  
XX 25-OCT-1996; 96WO-US17153.  
XX  
XX 25-OCT-1995; 95US-0548186.  
XX  
XX (ARRI-) ARRIS PHARM CORP.  
XX  
XX Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;  
XX  
XX WPI; 1997-259023/23.  
XX  
XX N-PSDB; AAT84526.  
XX  
XX Modified subtilisin inhibitor - having altered reactive site  
XX resulting in decreased immunogenicity, used in detergent  
XX composition(s)  
XX  
XX Example 2; Fig 11A; 69pp; English.  
XX  
XX This polypeptide comprises human pancreatic secretory trypsin  
XX inhibitor (PSTI), modified to include a Streptomyces subtilisin  
XX inhibitor (SSI) mutated reactive site. It is encoded by a DNA  
XX construct (AAT84526) produced by PCR amplification of overlapping  
XX synthetic oligonucleotides. Novel modified subtilisin inhibitors  
XX are claimed in which a human standard mechanism inhibitor, such as  
XX PSTI, has an altered reactive site so that its dissociation  
XX constant for inhibition of subtilisin is reduced by at least a  
XX factor of 100. Also claimed are: a nucleic acid encoding the  
XX subtilisin inhibitor; expression vectors; and transformed host  
XX cells. The novel subtilisin inhibitors combine the low  
XX allergenicity of human standard mechanism inhibitors, which are  
XX not specific for subtilisin, and the high affinity of non-human  
XX subtilisin inhibitors such as SSI. They form a complex with  
XX subtilisin, so as to avoid problems of allergenicity in detergent  
XX formulations, and can also be used in affinity purification and  
XX (diagnostic) quantification of subtilisin.

```

XX SQ Sequence 78 AA;
Query Match 17.9%; Score 76.5; DB 18; Length 78;
Best Local Similarity 35.4%; Pred. No. 0.092;
Matches 17; Conservative 5; Mismatches 21; Indels 5; Gaps 1;

Oy 35 DAKAAMTELKSCIDGLOPMHKAELVQLVQLGSDGAGTDYKDDDDK 82
Db 36 DGNTPNECVLCFEG-----RKRTSILIQKSGCGGGSDYKDDDDK 78

RESULT 15
AAB87666
ID AAB87666 standard; protein; 102 AA.
AC AAB87666;
XX 15-MAY-2001 (first entry)
DT Bovine mammary tissue derived protein #57.
DE Bovine; mammary gland; cancer; tumour; angiogenesis.
KW Bos taurus.
OS WO200114553-A1.
PN 01-MAR-2001.
PD 23-AUG-2000; 2000WO-N200166.
PF 23-AUG-1999; 99US-0150330.
PR (GENE-) GENESIS RES & DEV CORP LTD.
PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
XX Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;
XX WPI; 2001-226619/23.
XX New polypeptides and polynucleotides encoding the polypeptides, which
PT are expressed in bovine mammary gland tissue, useful for stimulating
PT mammary gland growth or function, or inducing differentiation of milk
PT producing cells.
XX Claim 11; Page 83; 97pp; English.
XX The present invention relates to proteins derived from bovine
CC mammary gland cells. The invention is useful for stimulating
CC bovine mammary gland cell growth and function, inhibiting the
CC growth of various mammary gland cancer cells, inhibiting
CC angiogenesis and vascularization of tumours, or modulating
CC the growth of blood vessels in a mammal.
XX SQ Sequence 102 AA;
Query Match 17.9%; Score 76.5; DB 22; Length 102;
Best Local Similarity 30.8%; Pred. No. 0.13;
Matches 20; Conservative 10; Mismatches 26; Indels 9; Gaps 1;

Oy 2 DNDFFMDFLQTLVGTPEELVGLKYNVNDKAAAMTELKSCIDGLOPMHKAELVKL 61
Db 30 DLTEFFY-----FPDLLYRLSLAKYNAPPEAFAKMEVKOCTDRFSVKNRLIITNI 80

Oy 62 LVQVL 66
Db 81 LGKIL 85

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:36:31 ; Search time 25.732 Seconds  
(without alignments)  
832.774 Million cell updates/sec

Title: US-09-768-826-47\_COPY\_2\_105  
Perfect score: 537  
Sequence: 1 KGRALLLVALTFCICRMA.....VOVLGSDGAGTDYKDDDK 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	87.9	95	4 Q8TD33	Q8td33 homo sapien
2	344.5	64.2	94	11 Q05702	Q05702 rattus ratt
3	90.5	16.9	96	11 Q8VD96	Q8vd96 mesocricetu
4	87.5	16.3	478	5 Q62571	Q62571 suberites d
5	76.5	14.2	504	11 Q9D008	Q9d008 mus musculu
6	76	14.2	2104	16 Q8XM24	Q8xm24 clostridium
7	74	13.8	457	5 Q9VAB8	Q9vab8 drosophila
8	73	13.6	90	6 Q9GK67	Q9gk67 oryctolagus
9	73	13.6	609	3 Q9UT83	Q9ut83 schizosacch
10	72.5	13.5	500	11 Q99J29	Q99j29 mus musculu
11	70.5	13.1	3071	12 Q8QXL2	Q8qx12 sorghum mos
12	70.5	13.1	3071	12 Q8QXL1	Q8qx11 sorghum mos
13	69.5	12.9	643	10 Q9FW93	Q9fw93 oryza sativ
14	69.5	12.9	2458	2 Q51827	Q51827 pseudomonas
15	69	12.8	540	16 Q9CFY4	Q9cfy4 lactococcus
16	68.5	12.8	2100	3 P87112	P87112 schizosacch

Q95247 homo sapien  
Q95277 homo sapien  
Q9uiq0 homo sapien  
Q81081 human parai  
Q8cy89 methanopyru  
Q19742 caenorhabdi  
Q8vxx7 arabidopsis  
Q8qzr1 mus musculu  
Q9lvq2 arabidopsis  
Q13450 saccharomyc  
Q80740 arabidopsis  
Q9hvl5 pseudomonas  
Q8ycg5 bruceella me  
Q8ryj9 oryza sativ  
Q85808 borrelia bu  
P94238 borrelia bu  
Q44978 borrelia bu  
Q95jy6 macaca fasc  
O57942 pyrococcus  
Q9xsw4 mustela vis  
Q62006 branchiosto  
Q86653 streptomyce  
Q35651 mus musculu  
P97790 mus musculu  
P97789 mus musculu  
Q97c56 thermoplasm  
Q9bps1 bombyx mori  
Q58122 pyrococcus  
Q9frn0 oryza sativ

#### ALIGNMENTS

##### RESULT 1

Q8TD33 PRELIMINARY; PRT; 95 AA.  
ID Q8TD33  
AC Q8TD33;  
DC 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Putative ligand binding protein RYD5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bingle C.D.;  
RT "Human RYD5, a new secretogloblin.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY026938; AAK08972.1;  
SQ SEQUENCE 95 AA; 10457 MW; F629AF06C96D2392 CRC64;

Query Match 87.9%; Score 472; DB 4; Length 95;  
Best Local Similarity 98.9%; Pred. No. 7.3e-45;  
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KGRALLLVALTFCICRMATGDNDEFFMDFLQTLVLTGTPTEELYEGTLGKYNVEDAKA 60  
Db 2 KGRALLLVALTFCICRMATGDNDEFFMDFLQTLVLTGTPTEELYEGTLGKYNVEDAKA 61  
Qy 61 AMTELKSCIDGLQPMHKAELVKLLVQLGSDGA 94  
Db 62 AMTELKSCRDGLQPMHKAELVKLLVQLGSDGA 95

##### RESULT 2

Q05702 PRELIMINARY; PRT; 94 AA.  
ID Q05702  
AC Q05702;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Potential ligand-binding protein.
GN RYD5.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=92007724; PubMed=1915264;
RA Dear T.N., Boehm T., Keverne E.B., Rabbitts T.H.;
RT "Novel genes for potential ligand-binding proteins in subregions of
RT the olfactory mucosa.";
RL EMBO J. 10:2813-2819(1991).
DR EMBL: X60661; CAA43068.1; -
DR InterPro: IPR000329; Uterogloblin.
DR SMART: SM00096; UTG; 1.
SQ SEQUENCE 94 AA; 10401 MW; 25A4BBB4977E247 CRC64;

Query Match 64.2%; Score 344.5; DB 11; Length 94;
Best Local Similarity 75.0%; Pred. No. 1.1e-30;
Matches 69; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

Qy 1 KGSRALLVLTFCICRMATGDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAKA 60
Db 2 KSSA-LLVLTVCICGLTRAEEDNEFFMDFLQTLVGTPEELYEGTLGKYNVEDAKA 60
Qy 61 AMTELKSCIDGLQPMHKAELVKLLVQLGSD 92
Db 61 ALTELKSCIDELQPVHKEQLVQLVQLDAQE 92

RESULT 3
Q8VD96 PRELIMINARY; PRT; 96 AA.
AC Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CC10 protein precursor.
GN CC10.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Gutierrez-Sagal R., Nieto A.;
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara
RT cell 10 kDa protein.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: L37041; AAL31349.1; -
DR InterPro: IPR003628; Uteroglobn_sub.
DR InterPro: IPR000329; Uterogloblin.
DR Pfam: PF01099; Uterogloblin; 1.
DR ProDom: PD012475; Uteroglobn_sub; 1.
DR SMART: SM00096; UTG; 1.
DR PROSITE: PS00404; UTEROGLOBIN_2; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.
FT SIGNAL.
SQ SEQUENCE 96 AA; 10509 MW; 5EB9CBDD46143389 CRC64;

Query Match 16.9%; Score 90.5; DB 11; Length 96;
Best Local Similarity 27.8%; Pred. No. 0.019;
Matches 25; Conservative 17; Mismatches 43; Indels 5; Gaps 2;

Qy 7 LLVLTFCICRMATGDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAKAAMELTK 66
Db 5 ITMAVVLVSVCCSSASSDTCPGFFQVLEFLFMGS-ESSYEALKFYFPGSLQDSGTQLK 63
```

```
Qy 67 SCIDGLQPMHKAELVKL---LVQVLGSD 92
Db 64 KLVDTLPQKTRNINMKLSEILTLSPLCNQD 93

RESULT 4
O62571 PRELIMINARY; PRT; 478 AA.
AC O62571;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Stress-responsive protein kinase (PRKSD).
OS Suberites domuncula (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Hadromerida; Suberitidae; Suberites.
OX NCBI_TaxID=55567;
RN [1]
RP SEQUENCE FROM N.A.
RC Mueller W.E.G.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96394688; PubMed=8798342;
RA Kruse M., Gamulin V., Cetkovic H., Pancer Z., Mueller I.M.,
RA Mueller W.E.G.;
RT "Molecular evolution of the Metazoan protein kinase C multigene
RT family.";
RL J. Mol. Evol. 43:374-383(1996).
DR EMBL: Y13101; CAA73555.1; -
DR HSP; P24941; LHCL.
DR InterPro: IPR00719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 478 AA; 54098 MW; 6638FF5FD88D6332 CRC64;

Query Match 16.3%; Score 87.5; DB 5; Length 478;
Best Local Similarity 31.0%; Pred. No. 0.31;
Matches 27; Conservative 18; Mismatches 25; Indels 17; Gaps 4;

Qy 21 TGEONDEF---FMDFLQTLVGTPEELYEGT-LGKYNVEDAKAAMELTKSCIDGLQPMH 76
Db 238 TLKDTDFSNFSDFISRCVLKPNPEERMSATALLQHKFSKAK-PVAVLK----- 286

Qy 77 KAEVLKLLVQVLGSDGAGTDYKDDDD 103
Db 287 --ELIQDAMRILEEEGSGSDEDDND 311

RESULT 5
Q9D008 PRELIMINARY; PRT; 504 AA.
AC Q9D008;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Signal recognition particle 54 kDa.
GN SRP54.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
```

[illegible]



```
DR SMART: SMO0382; AAA: 1.
DR PROSITE; PS00300; SRP54; UNKNOWN_1.
SQ SEQUENCE 500 AA; 55492 MW; C957B6F76C6A224B CRC64;

Query Match 13.5%; Score 72.5; DB 11; Length 500;
Best Local Similarity 32.4%; Pred. No. 15;
Matches 36; Conservative 14; Mismatches 42; Indels 19; Gaps 6;

Qy 1 KGSRALLVALLTFLCICRMATGENDDEF---FMDFLQTLV-VGTPPEELYEGTLGKYNNV 55
   || || || || || || || || || || || || || || || || || || || || ||
Db 255 KGGGALSAVATKSPFIIGTGEHIDNEPEKTOPFFISKILCMGDIEGLID-----KVN 308
   || || || || || || || || || || || || || || || || || || || || ||

Qy 56 E-----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSDQAGTDY 98
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 ELKLDNKKALIEKWKHGQFTLRDMYEQFQNTIMKMGPFQSOILGMPGFGTDF 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
Q8QXL2
ID O8QXL2 PRELIMINARY; PRT; 3071 AA.
AC O8QXL2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein.
OS Sorghum mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=32619;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XIAOSHAN;
RA Chen J., Chen J.P., Adams M.J.;
RT "Characterisation of potyviruses from sugarcane and maize in China.";
RL Arch. Virol. 147:0-0(2002).
DR EMBL; AJ310197; CAC84437.1; -
FT CHAIN 2 233 P1 PROTEIN.
FT CHAIN 234 693 HC-PRO PROTEIN.
FT CHAIN 694 1040 P3 PROTEIN.
FT CHAIN 1041 1107 6K1 PROTEIN.
FT CHAIN 1108 1745 C1 PROTEIN.
FT CHAIN 1746 1798 6K2 PROTEIN.
FT CHAIN 1799 1987 NIA-VPG PROTEIN.
FT CHAIN 1988 2229 NIA-PRO PROTEIN.
FT CHAIN 2230 2750 NIB PROTEIN.
FT CHAIN 2751 3071 COAT PROTEIN.
SQ SEQUENCE 3071 AA; 349970 MW; 2FD59A0668DE2FB7 CRC64;

Query Match 13.1%; Score 70.5; DB 12; Length 3071;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 22; Conservative 10; Mismatches 17; Indels 17; Gaps 3;

Qy 21 TCEDNDEEFMDFLQT-----LLVGTPEELYEGTLGKYNNVEDAKAAMTELKSCIDGLQPMH 76
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2363 TGKKKD-YFMDFSQDTEKILKQSCERLYEGQLGVWNGSLKAE-----IRPIE 2409
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 77 KAEVLK 82
   | | | |
Db 2410 KTEANK 2415

RESULT 12
Q8QXL1
ID O8QXL1 PRELIMINARY; PRT; 3071 AA.
AC O8QXL1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Polyprotein.
OS Sorghum mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=32619;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YUHAN;
RA Chen J., Chen J.P., Adams M.J.;
RT "Characterisation of potyviruses from sugarcane and maize in China.";
RL Arch. Virol. 147:0-0(2002).
DR EMBL; AJ310198; CAC84438.1; -
FT CHAIN 2 233 P1 PROTEIN.
FT CHAIN 234 693 HC-PRO PROTEIN.
FT CHAIN 694 1040 P3 PROTEIN.
FT CHAIN 1041 1107 6K1 PROTEIN.
FT CHAIN 1108 1745 C1 PROTEIN.
FT CHAIN 1746 1798 6K2 PROTEIN.
FT CHAIN 1799 1987 NIA-VPG PROTEIN.
FT CHAIN 1988 2229 NIA-PRO PROTEIN.
FT CHAIN 2230 2750 NIB PROTEIN.
FT CHAIN 2751 3071 COAT PROTEIN.
SQ SEQUENCE 3071 AA; 349537 MW; E3D326DE1B9ACB05 CRC64;

Query Match 13.1%; Score 70.5; DB 12; Length 3071;
Best Local Similarity 33.3%; Pred. No. 2.5e+02;
Matches 22; Conservative 10; Mismatches 17; Indels 17; Gaps 3;

Qy 21 TCEDNDEEFMDFLQT-----LLVGTPEELYEGTLGKYNNVEDAKAAMTELKSCIDGLQPMH 76
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2363 TGKKKD-YFMDFSQDTEKILKQSCERLYEGQLGVWNGSLKAE-----IRPIE 2409
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 77 KAEVLK 82
   | | | |
Db 2410 KTEANK 2415

RESULT 13
Q9FW93
ID Q9FW93 PRELIMINARY; PRT; 643 AA.
AC Q9FW93;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative cytochrome P450.
GN OSJNBAA0026L12.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBAA0026L12 genomic sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC068924; AAG13498.1; -
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 643 AA; 70706 MW; C695CA5DF1A36463 CRC64;

Query Match 12.9%; Score 69.5; DB 10; Length 643;
Best Local Similarity 22.4%; Pred. No. 44;
Matches 33; Conservative 19; Mismatches 44; Indels 51; Gaps 6;

Qy 2 GSRALLVALLTFLCICRMATGEDNDEFFMDFLQTLVGTGP---EELYE-----GTLGKY 52
   || || || || || || || || || || || || || || || || || || || || ||
Db 296 GNRATDDELMLCSEFLVAGTDSAAALQWIMAEILVRNPISQSKLYERIKSKTGGGGH 355
   || || || || || || || || || || || || || || || || || || || || ||

Qy 53 NVNEDAKAAMTELKSCI-DGLQ---PMH-----KAELVKLLVQVLG 89
```

DB 356 EVSEEDVHDPYLKAVLEGLRKHPPAHMLPHKAAEDMDVGGYLPKGTIVNFWAENG 415  
QY 90 SQ-----DAGTGD 97  
DB 416 RDEKEWKPWFMPERFLPGDGEGVD 442  
RESULT 14  
O51827 PRELIMINARY;  
AC O51827; PRT: 2458 AA.  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Polyketide synthase type I.  
GN PLTB  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PF-5;  
RX MEDLINE=98094250; PubMed=9434161;  
RA Nowak-Thompson B., Gould S.J., Loper J.E.;  
RT "Identification and sequence analysis of the genes encoding a  
polyketide synthase required for pyoluteorin biosynthesis in  
Pseudomonas fluorescens Pf-5";  
RL Gene 204:17-24(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PF-5;  
RX MEDLINE=99194726; PubMed=10094695;  
RA Nowak-Thompson B., Chaney N., Wing J.S., Gould S.J., Loper J.E.;  
RT "Characterization of the pyoluteorin biosynthetic gene cluster of  
Pseudomonas fluorescens Pf-5";  
RL J. Bacteriol. 181:2166-2174(1999).  
DR EMBL; AF081920; AAC38075.1;  
DR InterPro; IPR001227; Ac transferase.  
DR InterPro; IPR000794; Ketoacyl-synt.  
DR InterPro; IPR003880; Ppantme\_attach.  
DR InterPro; IPR002155; Thiolase.  
DR Pfam; PF00698; Acyl\_transf; 1.  
DR Pfam; PF0109; ketoacyl-synt; 2.  
DR Pfam; PF02801; ketoacyl-synt\_C; 2.  
DR Pfam; PF00550; pp-binding; 2.  
DR PROSITE; PS00075; ACP\_DOMAIN; 2.  
DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 2.  
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
DR PROSITE; PS00098; THIOLASE\_1; UNKNOWN\_1.  
KW Phosphopantetheine; Transferase.  
SQ SEQUENCE 2458 AA; 262676 MW; AE756080AEIA5FB1 CRC64;

Query Match 12.9%; Score 69.5; DB 2; Length 2458;  
Best Local Similarity 32.1%; Pred. No. 2.4e+02;  
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;

QY 43 ELYEGTLGKYNVEDAKAMTELKSCIDGLQPMHKAELVLLVQLVLSQDGAG 95  
DB 2239 QLLHLTQSELEQDA-AARTAMQG--SGLOPLQRSQIVQAIARVLGGGQCG 2288

RESULT 15  
O9CFY4 PRELIMINARY;  
AC O9CFY4; PRT: 540 AA.  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Fibronectin-binding protein.  
GN YNGB OR L11327  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OX Streptococcaceae; Lactococcus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=1137471;  
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,  
RA Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
lactis ssp. lactis IL1403";  
RL Genome Res. 11:731-753(2001).  
DR EMBL; AE006365; AAK05425.1;  
KW Complete proteome.  
SQ SEQUENCE 540 AA; 61285 MW; 4DAC3CCCBFCF72867 CRC64;

Query Match 12.8%; Score 69; DB 16; Length 540;  
Best Local Similarity 30.6%; Pred. No. 41;  
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 24 DNDEFEM---DFLOTLLVGTPEELYEGTLGKYNVEDAKAMTELKSCIDGLQPMHKA 79  
DB 294 DNAELFROKGEILLTNFLNQVNDKTSVLTENYTTNEPIEIALNPALSPVQNAQRYFHR 353  
QY 80 LVKLLVQVLGSQ 91  
DB 354 KLKQAVKFLGEQ 365

Search completed: April 28, 2003, 14:43:30  
Job time : 26.732 secs